DIFFERENTIAL GENE EXPRESSION IN THE CULM OF SUGARCANE DURING DEVELOPMENT, WITH SPECIAL EMPHASIS ON THE STORAGE PARENCHYMA CELLS

by

Omeswaree Rogbeer

Thesis submitted in partial fulfilment of the academic requirements for the degree of

Master of Science

in

Plant Biotechnology

University of Stellenbosch

Supervisor: Prof. F.C. Botha

Co-supervisor: Dr. S. Groenewald

Stellenbosch

December 2002

DECLARATION

I, the underdesigned, hereby declare that the work contained in this thesis is my own original work and I have not previously in its entirety or in part submitted it at any university for a degree.

Date: 21 August, 2002

ABSTRACT

For the expression of transgenes in plant cells, appropriate promoter sequences have to be introduced upstream of the gene to ensure efficient transcription. While to date the maize ubiquitin (*Ubi1*) promoter has been the most effective transgene promoter for sugarcane, there is a high demand for tissue and stage specific promoters for localised transgene expression in the mature culm. The present study sought to characterise genes preferentially expressed in the core and peripheral tissues of the mature culm, which can further be used as research tools for specific promoter isolation.

cDNA expression arrays containing 3840 clones from a late stage cDNA library representative of the core and peripheral tissues of the mature culm were prepared. The cDNA expression arrays were then differentially screened in independent hybridisation experiments with radioactively-labeled cDNA representations of core and peripheral tissues of internode 7, and peripheral tissues of internode 10. Comparison of the expression profiles of the arrayed cDNA targets in the three probes led to the identification of 60 tissue-specific, 17 stage-specific and 50 selectively expressed cDNAs within the mature sugarcane culm.

ESTs of 33 chosen selectively expressed cDNAs with a relatively stronger pattern of expression in the core than in the peripheral tissues revealed sequence homology to a diverse collection of genes in the mature culm. These included genes associated with general cellular metabolism such as protein synthesis, protein modification and structural protein. Also identified were stress-responsive genes. The putative translational products of some of these clones had homologs that are involved in cell-wall structure in other species. These included the jacalin homolog, a lectin, hydroxyproline rich glycoprotein and structured polyprotein C. Many of the cDNAs thought to be involved in cell wall structure or stress related responses also accumulate in a developmental manner in other plants. These may indicate that specific mature culm mRNAs accumulate in response to stresses such as rapid cell expansion or as part of the late developmental program. An unexpected

observation was that only one gene associated with sucrose metabolism was identified, namely sucrose synthase. These results confirmed that culm maturation was not controlled by sucrose metabolism despite its distinct physiological characteristic of storing high levels of sugars.

ESTs analysis further revealed that sequence homology was not obtained for all the cDNAs exhibiting stage and tissue specific expression in the core and peripheral tissues of the mature culm. These could represent novel genes not only from sugarcane but all plants.

Northern analysis demonstrated that 9 putatively identified selectively expressed genes tested so far accumulated specifically in the core and peripheral tissues of the mature culm. No expression was detected in root, leaf, leafroll and internode 3. However, their selective expression in a single internode as observed on the arrays (i.e hybridisation signal intensity being higher in the core than in the peripheral tissue) was not detected on the northern blots. These showed that cDNA expression arrays were not a high-capacity gene expression assay since they were prone to false expression analysis. The validity of results obtained through array screening should always be verified in an independent manner, preferably by the northern hybridisation analysis.

Hence, the present study shows that the combination of differential screening, northern blot and DNA sequence analysis permits the rapid characterisation of differentially expressed genes in the core and peripheral tissues of the mature sugarcane culm. These can further be used as research tools for mature culm – specific promoter isolation in the sugarcane.

OPSOMMING

Die doeltreffende uitdrukking van transgene in plantselle is afhanklik van 'n gepaste promotorvolgorde wat stroomop van die geen ingevoeg word. Die *Ubi1*-promotor van mielies was tot dusver die doeltreffendste transgeenpromotor in suikerriet, maar daar is 'n groot behoefte aan promotors wat weefsel- en ontwikkelingstadium-spesifieke geenuitdrukking kan beheer. Hierdie studie het op die isolering en karakterisering van gene wat selektief in die kern- of periferale stingelweefsel van suikerriet uitgedruk word, gefokus. Hierdie gene sal verder benut kan word om promotors te isoleer.

cDNA uitdrukkingsreekse ("expression arrays") van 'n volwasse stingel cDNA biblioteek is voorberei. Hierdie reekse, wat 3840 klone bevat het, is in onafhanklike hibridiseringseksperimente met radioaktiefgemerkte cDNA van onderskeidelik kern- en periferale stingelweefsel van lit 7 en periferale stingelweefsel van lit 10 afgetas. 'n Vergelyking van die uitdrukkingsprofiele van die cDNA teikens in dié drie peilergroepe het tot die identifisering van 60 weefsel-spesifieke-, 17 ontwikkelingstadium-spesifieke- en 50 selektief uitgedrukte cDNAs in die volwasse suikerrietstingel gelei.

Uitdrukkingsvolgordemerkers ("ESTs") van 33 geselekteerde cDNAs wat in hoër vlakke in die kern uitgedruk is, se volgordes toon homologie aan 'n wye verskeidenheid gene in die volwasse stingel. Hierdie groep sluit gene in wat met algemene sellulêre metabolisme soos proteïensintese, proteïenmodifisering en strukturele proteïene geassosieer is. Spanningsverwante gene is ook hier geïdentifiseer. Die transleringsprodukte van sommige klone het homoloë wat by selwandstruktuur in ander spesies betrokke is, soos die jacalin-homoloog, 'n lektien, hidroksiprolien-ryke glikoproteïen en gestruktureerde poliproteïen C. 'n Wye verskeidenheid cDNAs wat by selwandstruktuur of spanningsverwante reaksies betrokke is, akkumuleer ook in 'n ontwikkelingsafhanklike wyse in ander plante. Dit mag 'n aanduiding wees dat spesifieke mRNAs in die volwasse stingel in reaksie op spanning wat met vinnige seluitsetting gepaardgaan, versamel. Slegs een geen wat met sukrose metabolisme geassosieer is, nl. sukrosesintase, is in hierdie studie geïdentifiseer. Hierdie onverwagte waarneming het bevestig dat, ondanks suikerriet se kenmerkende vermoë om hoë konsentrasies suiker te berg, stingelveroudering nie net met sukrose metabolisme geassosieer kan word nie. Nie al die cDNA-fragmente wat geïsoleer is,

het homologie aan ander gene in die internasionale databasisse getoon nie, wat moontlik kan aandui dat nuwe gene suksesvol geïsoleer is.

Nege ontwikkelingstadium-spesifieke gene wat slegs in die volwasse stingelweefsels uitgedruk word, is dmv noordelike oordraganalises geïdentifiseer. Geen transkripte van hierdie gene is in die wortels, blaarrol, blare of jong stingel waargeneem nie. Die weefselspesifisiteit wat met die uitdrukkingsreekse waargeneem is, kon nie mbv noordelike orrdraganalises bevestig word nie. Dit mag 'n aanduiding wees dat die uitdrukkingsreekse vals positiewe resultate kan oplewer en dit is raadsaam om voortaan altyd die verkrygde profiele met ander, meer sensitiewe tegnieke, te bevestig.

Die studie het aangetoon dat 'n kombinasie van differensiële aftasting, noordelike oordraganalise en DNA-volgordebepaling gebruik kan word om gene wat differensieel uitgedruk word in die volwasse suikerrietstingel, te identifiseer. Hierdie geenfragmente kan nou vir promotorisoleringsdoeleindes aangewend word.

ACKNOWLEDGEMENTS

My sincere thanks and appreciation go to my supervisor, Prof. F.C. Botha (Director, Institute for Plant Biotechnology), for his valuable support during the course of the study. This support came in the form of both academic and personal advice when it was mostly needed. His enthusiasm, optimism, encouragement and endless patience have made this study an enriching experience.

My special thanks and appreciation are also extended to the following:

° My co-supervisor, Dr. Sarita Groenewald, for her expertise in several areas, valuable suggestions and support during the course of this investigation, critical proof reading of the manuscript and above all, her friendship.

° Dr. Bernard Potier for his assistance and expert advice in several areas during the course of this study.

° Dr. Barbara Huckett and Prof. Robert Birch for their advice and encouragement during their visits at the IPB.

° Carl van Heerden for his assistance with sequencing.

° The South African Sugar Industry for granting me the financial support to carry out this study.

° All my colleagues and friends at the Institute for Plant Biotechnology for advice and encouragement, whenever it was needed.

° My parents and special brother Saleel for their love, support and patience during this study.

° A special thanks go to the Chetty family, most specially to Kuben for their support and encouragement over the years

A DEDICATION

My thesis is dedicated to my supervisor Prof. F.C. Botha, my co-supervisor Dr. Sarita Groenewald, my family and Kuben Chetty for their love, support, interest and encouragement.

" Now to him who is able to do immeasurably more than all we ask or imagine, according to his power that is at work within us, to Him be the glory…"

EPH. 3:20,21

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ABBREVIATIONS

35S	cauliflower mosaic virus *35S* promoter
A	absorbance
ABA	abscisic acid
ABRE	abscisic acid response element
Act1	rice actin-1 promoter
Adh1	maize alcohol dehydrogenase-1 promoter
AFLP	amplified fragment length polymorphism
albD	albidicin detoxification
AMV	Avian Myeloblastosis Virus
ATP	adenosine 5' - triphosphate
azac	azacytidine
bar	bialaphos resistance gene
BLAST	basic local alignment search tool
bp	base pair
CaMV	cauliflower mosaic virus
Cat	chloramphenicol acetyl transferase gene
cDNA	complementary deoxyribonucleic acid
cpm	counts per minute
dATP	deoxyadenine 5' – triphosphate
dbEST	expressed sequence tag database
dCTP	deoxycytosine 5' – triphosphate
DDRT-PCR	differential display reverse transcription polymerase chain reaction
DEPC	diethylpyrocarbonate
DNA	deoxyribonucleic acid
dNTP	deoxynucleoside triphosphate
DRE	dehydration response element
ds	double stranded
DSE	downstream element
EDTA	ethylenediaminetetraacetate
EST	expressed sequence tag
FSD	family specific domain
g	gravitational accelaration (9.806 m s^{-1})

GUS	β - glucuronidase (β-D-*Glucuronide glucuronosohydrolase*), EC 3.2.1.31
HSE	heat shock element
HSPs	heat shock proteins
IPTG	isopropyl-β-D-thiogalactoside
kb	kilobase
LB	Luria Bertani
MAST	magnet–assisted subtraction technique
MCB	Mlu cell-cycle box
MOPS	3-[*N*-morpholino]-propanesulphonic acid
mRNA	messenger ribonucleic acid
nos	nopaline synthase gene
ocs	octapine synthase gene
onc	oncogenes
PAM	Point Acceptable Mutation
PEG-6000	polyethylene glycol 6000
PVPP	polyvinyl polypyrrolidone
PTGS	post transcriptional gene silencing
rbcS	ribulose-1,5-bisphosphate carboxylase-oxygenase gene
RC4D	RFLP – coupled domain directed differential display
RFLP	restriction fragment length polymorphism
RNA	ribonucleic acid
RNases	ribonucleases
rRNA	ribosomal ribonucleic acid
RT	reverse transcription
RT-PCR	reverse transcription polymerase chain reaction
RYMV	rice yellow mottle virus
SAGE	serial analysis of gene expression
SDS	sodium dodecyl sulphate
SH	subtractive hybridisation
smHSPs	small heat shock proteins
SNF1	sucrose non-fermenting 1
SPS	sucrose-phosphate synthase (Sucrosephosphate-UDP glucosyltransferase), EC 2.4.1.14

SSU	small subunit
SuSy	sucrose synthase (Sucrose-UDP glucosyltransferase), EC 2.4.1.14
tRNA	transfer ribonucleic acid
Tris	Tris(hydroxymethyl)-aminomethane
Ubi1	maize polyubiquitin-1 promoter
UV	ultraviolet
Vvht1	*Vitis vinifera* hexose transporter 1
v/v	Volume/volume
w/v	Weight/volume

CHAPTER 1
Introduction

Genetic manipulation of plants over the past two decades has led to significant progress in the field of plant biotechnology. The capacity to introduce and express diverse foreign genes in plants, firstly described by Ledoux and Huart (1969), has been extended to over 120 species and in at least 38 families (Dale, 1995). Successes include a few major economic crops, vegetables, ornamental, medicinal, fruit, tree and pasture plants. More than 3000 field trials of plants, genetically modified for various economic traits, are in progress or have been completed in at least 31 countries (Dale, 1995).

This spectacular progress was mainly due to the development of requisite molecular tools for the manipulation of primary metabolic pathways in plants (Sonnewald and Willmitzer, 1992). These were (1) suitable transformation methods and tissue culture procedures to generate a sufficiently high number of regenerants, (2) isolation of suitable genes from prokaryotic and eukaryotic sources, (3) identification of suitable gene promoter elements to direct strong tissue/organ- and cell-specific expression, and (4) identification of suitable targeting sequences to direct the gene product to the appropriate subcellular location (Christopher and Campbell, 2001). Despite these developments and their applications in dicot plant species, ranging from the model plant systems *Arabidopsis* and *Nicotiana* to agronomically significant crops such as potato, tomato and cotton (Christopher and Campbell, 2001), molecular tools for the transformation of sugarcane and other monocots are lagging far behind.

However, significant progress in monocot transformation has been made and since 1989, reliable reports on the transformation of monocotyledonous plants such as rice (Shimamoto *et al.*, 1989), maize (Gordon-kamm *et al.*, 1990), wheat (Vasil *et al.*, 1992), barley (Wan and Lemaux, 1994) and sugarcane (Bower and Birch, 1992) have been published. In some of the transformed plants, the foreign genetic material was simply no longer expressed. This phenomenon known as gene silencing became the cause of major concern in transgenic plant programmes (Finnegan and McElroy, 1994). Since promoters

offer a fundamental role in expressing the coding regions of given genes that are included in DNA constructs for direct transformation, a great deal of research interest is geared at isolating and studying promoters.

Currently, the promoters commonly used in monocot transformation include promoters of dicot origin, for example the *35S* promoter from the cauliflower mosaic virus (CaMV) (Elliott *et al.,* 1998) as well as constitutive promoters of monocot origin, namely Emu (based on a truncated *Adh1* promoter from maize) (Rathus *et al.,* 1993), *Act1* (rice actin-1) (Grof *et al.,* 1996a) and the *Ubi1* (maize polyubiquitin-1 promoter) (Gallo-Meagher and Irvine, 1993). However, while constitutive promoters used in the sugarcane transformation may be appropriate for the expression of certain types of transgenes, there is a growing need for promoters that could regulate a desired pattern of gene expression in specific tissues and stages of the plant. For example, tissue and stage specific promoters can be used to alter sucrose metabolism in the sugarcane culm for increased sucrose yield.

In addition, these promoters can be used to specifically express resistance genes in the sugarcane culm, the target for pests and diseases such as the stalk borer *Eldana saccharina* Walker, smut and rust, which cost the sugar industry a huge amount of money per season through direct damage. Furthermore, use of tissue and stage specific promoters in the sugarcane culm would also relieve the plant from the strain of transgene expression throughout plant growth and development. Lastly, the development of a promoter, being effectively an invention, is eligible for patent protection. Patents have already been issued on most established promoters. Hence, this has necessitated a renewed effort to isolate novel regulatory elements from the sugarcane.

One possibility to obtain promoters directing such tailored patterns of expression is to look for genes already expressed in the desired pattern in the sugarcane, and then to isolate the corresponding promoters from the sugarcane genome (Hansom *et al.*, 2001). Numerous genes and associated promoters, which exhibit a wide range of tissue and/or developmental

expression patterns, have been described. Examples include genes, which are specifically expressed in pollen (Custers *et al.*, 1997), flowers, fruit (Coupe and Deikman, 1997), embryos, cotyledons, endosperm and phloem (Clark *et al.*, 1997; Yang and Russel, 1990).

In contrast, little is known about such genes in the sugarcane culm. However, changes in morphological features of the sugarcane culm, such as an increase in the number of vascular bundles but a decrease in size from the core to the peripheral tissues, an increase in lignification and suberisation down the mature stalk in parallel with increased sucrose storage, have been observed through histochemichal and sugar assays (Jacobsen *et al.,* 1992). From these observations, it is hypothesised that genes are differentially expressed in the different tissue-types of the sugarcane culm and at different developmental stages. Since sucrose is accumulated in the culm and some pathogens gain entry through the buds, the culm has become one of the major target areas for transgene expression in the sugarcane. It is therefore the focus area in this study to identify differentially expressed genes specific to the core and peripheral tissues of the mature sugarcane culm, which would be used as research tools for the isolation of specific promoters.

Genes are differentially regulated, leading to different mRNA transcript levels, between different tissues and at different developmental stages of a plant, under the control of specific regulatory elements present within the minimal promoter region (Fang *et al.*, 1989). Specific regulatory elements are genomic sequence motifs that are over-represented in the genomic DNA in the vicinity of similarly behaving genes (Fang *et al.*, 1989). They are also referred to as *cis* acting regulatory elements. Transcriptional activation of genes controlled by typical *cis* acting DNA sequences is further discussed in the literature review, which does not summarise all aspects of plant promoters but is rather focused on the characterisation of regulatory domains conferring specific promoter qualities. In addition, strategies previously employed for specific gene isolation are also reported (Chapter 2).

In an attempt to identify differentially expressed genes specific to the core and peripheral tissues of the mature sugarcane culm in this study, the primary approach was to perform surveys of transcribed gene expression patterns in the tissues mentioned above. To investigate gene expression based entirely on gene transcriptional level, obtaining good quality total RNA with intact mRNA was crucial for the development of a specific late stage cDNA library, representative of the genes expressed in the two tissue types of the mature culm and for northern analysis. Hence, emphasis was placed on the choice of an appropriate RNA extraction method applicable to the mature stem tissues of the sugarcane.

As a secondary approach to this investigation, cDNA expression arrays representative of the late stage cDNA library were differentially screened to identify classes of highly expressed, differentially regulated genes specific to the two tissue types and stages of the mature sugarcane culm. The advantage of using a differential screen based only on differential mRNA levels, was that it required no specific biochemical information relating to the genes to be investigated (Desprez *et al.*, 1998).

However, to limit the results of a screen to relevant transcripts, the chosen tissues from which mRNA samples were prepared for hybridisation probe construction, were the core and peripheral tissues of the mature culm. The abundance of differential transcripts in hybridisation probes were maximised by preparing these from one of the cDNA preparations used to construct the late developmental stage cDNA library and from tissues at temporal stages of the sugarcane mature culm from which maximal expression of the desired genes were expected. DNA sequence analysis was then performed for the characterisation of identified differentially expressed genes and northern blots to further confirm their gene expression pattern. In chapter 3, experimental details of the research technique is described, followed by a report of the data obtained throughout the study (Chapter 4). The significant contribution of this study to the South African sugar research is then discussed in chapter 5.

CHAPTER 2
Literature Review

2.1 Introduction

The long-term goal of this study is the isolation of specific promoters in the sugarcane culm at late developmental stages, using differentially expressed genes as a research tool. A comprehensive review on promoters and their use in transformation of dicotyledonous plants has been reported over the past decade (Galun and Breiman, 1997). This chapter is not an attempt to cover all aspects of plant promoters but is rather focused on the characterisation of regulatory domains conferring specific promoter qualities.

It also reports on the development of genetic constructs, which has paved the way for the successful transformation of monocotyledonous plants such as maize (Gordon-Kamm *et al.*, 1990), rice (Shimamoto *et al.*, 1989) and wheat (Vasil *et al.*, 1992), over the past decade. While to date the maize polyubiquitin (*Ubi1*) promoter (Christenson *et al.,* 1992) has been the most effective transgene promoter for sugarcane, ongoing research is seeking to identify and characterise native, organ-specific promoters from sugarcane (Christopher and Campbell, 2001). An improved control of transgene expression through specific promoters and genomic integration, are likely to be key developments that are necessary to fulfill the intended agronomic function of each new gene in the sugarcane plant (Hansom *et al.,* 2001).

Henceforth, this chapter provides an overview of numerous research techniques used for the isolation and identification of specific genes.

2.2 Detailed Analysis of a Minimal Promoter

In DNA constructs, there are three important elements, namely, the promoter, coding sequence and terminator sequences. The promoter element determines the level of expression, developmental timing and pattern of tissue expression in transgenic plants. Adjacent to the 3' end of the transgene, a terminator DNA signal is incorporated into the construct and it contains consensus signals such as AATAA and a less well defined downstream element, (DSE) (Sullivan and Green, 1993). Both the polyadenylation (AATAA) and the DSE in synthetic terminal regions are involved in the correct maturation of the processed mRNA.

In plants, promoter elements are not only characterised by the system *cis* acting regulatory elements but a defined leader sequence which is transcribed but not translated. The leader sequence contains several elements and boxes such as the m^7 GPPPG "cap", which are essential for the efficient translation of the coding sequence (Sullivan and Green, 1993).

In most promoters, the proximal region contains a TATA (T/A) A (T/A) box, an AT rich sequence located 21 to 35 nucleotides upstream from the transcription start site (Tjian and Maniatis, 1994). This region is surrounded by GC rich sequences and interacts with a *trans* acting TATA-binding protein, which comprises the transcription factor TFIID. The latter forms a stable initiation complex that is necessary for accurate transcription of eukaryotic genes by RNA polymerase II (Rosenthal, 1987).

In some genes, the intact TATA box also determines cell or organ specific expression (McCormick *et al.*, 1994). However, studies based on the deletion of the TATA box resulted in decreased but not inhibited promoter activity (An and Kim, 1993). In such cases, the initiator *cis* element, which is a less defined consensus sequence, initiated transcription via binding transcription factors for the placement of the start site (Smale and Baltimore, 1989). Although the requirement

of the TATA box have been confirmed by analysis of promoter mutations in transgenic plants (Morelli *et al.*, 1985; Zhu *et al.*, 1993), the functional architecture of the proximal regions of a promoter have been less analysed.

Distal *cis* regulatory elements have been characterised that modulate or stabilise the formation of the initiation complex at the TATA box (Mukherjee *et al.*, 1995). A hexamer sequence, TGACGT, located within a few hundred nucleotides upstream from the transcription start site is crucial in the promoter activity of the cauliflower mosaic virus (CaMV) *35S*, octapine synthase (*ocs*) and nopaline synthase (*nos*) promoters (An and Kim, 1993). The hexamer motifs are often found as repeats separated by six to eight nucleotides.

Enhancer elements are another group of *cis* regulatory sequences that are involved in increasing the level of transcription. Their conserved motif in plants GTGTGG (T/C)(T/C)A(A/T)TA(T/A)G is bound by a region of 100-200 base pairs (bp) in length and can function from a distance of several thousand base pairs from either orientation of the respective gene (Lefebvre and Gellatly, 1997). The CCAAT box sequence is similarly functional from both 5' and 3' ends of the coding sequence but is mostly located at -80 in several promoters (Benfey and Chua, 1990). Thus, identified and unknown *cis* acting elements form the combinatorial nature or control of promoter activity regulating gene expression, classifying promoters in 2 categories:

i) promoters regulating constitutive expression, and
ii) promoters driving heterologous or specific expression.

2.3 Promoters for Monocotyledonous Plants

Over the past decades, both constitutive and non-specific promoters have been efficient in the transformation of monocotyledonous plants such as maize (Gordon-Kamm *et al.*, 1990), rice (Shimamoto *et al.*, 1989), wheat (Vasil *et al.*, 1992) and barley (Wan and Lemaux, 1994), once thought to be recalcitrant plants. In

transformation procedures where constitutive promoters are employed, expression of desired genes can be screened from all regenerating tissues of the transgenic plant. Little variation of mRNA abundance or in its translational product is observed during development, in different organs or upon application of various endogenous or environmental stimuli (Kuhlemeier *et al.*, 1987). Examples of such promoters are the nuclear gene promoter of the β – subunit of the mitochondrial ATPase complex (Boutry and Chua, 1985) and the rice actin-1 (*Act1*) promoter (Zhang *et al.*, 1991). A non-specific promoter such as the CaMV *35S*, differs slightly from the constitutive promoters in that they drive expression of transgenes in most but not all tissues of the plant.

In the early transformation studies with rice and maize (Vasil *et al.*, 1994), the CaMV *35S* promoter was used to activate the selective and/or reporter genes. Decreased strength or activity of the promoter up to 100 fold was observed, compared to its efficient regulatory role in dicotyledonous plants (Vasil *et al.*, 1994). Hence, in a strategy to increase levels of gene expression in cereal crops, genetic constructs were made by duplicating the CaMV *35S* promoter sequences (Vasil *et al.*, 1994). Advanced research led to the observation that cereals frequently have an intron in the transcribed but not translated 5' leader sequence of their genes. This finding provided ideas for designing new genetic constructs including such promoters or their introns in combination with the CaMV *35S* promoter (Fromm *et al.*, 1990). This concept was followed with the cereal alcohol dehydrogenase1 (*Adh1*) gene (Callis *et al.,* 1987; Kyozuka *et al.,* 1990). In several transformation studies, the whole *Adh1* or its intron, in combination with the *35S* promoter were used to transform cereal crop plants (e.g in maize: Fromm *et al.,* 1990; in wheat: Vasil *et al.,* 1992). Omirulleh *et al.* (1993) found that a promoter that was composed of a double enhancer element from the *35S* promoter fused to a truncated promoter of the wheat α-amylase gene, caused strong expression of a reporter gene in transgenic maize.

A similar approach to integrate the first intron of the rice actin-1 (*Act1*) gene was also followed in cereal transformation (Zhang *et al.*, 1991). The rice *Act1* promoter and its intron were found to be even more potent than the previously mentioned cereal promoters and their respective introns (Zhang *et al.*, 1991). The *Act1* promoter seemed to be at a similar level of potency as the monocot pEmu promoter (Last *et al.,* 1991). The latter promoter is a recombinant promoter containing a truncated maize *Adh1* promoter with multiple copies of the anaerobic response element from maize *Adh1* gene and *ocs* elements from the octapine synthase gene of *Agrobacterium tumefaciens* (Last *et al.,* 1991; Rathus *et al.,* 1993).

Finally, the promoter most effective in expressing stable transgene expression during transformation of rice, wheat, barley and sugarcane is the polyubiquitin-1 (*Ubi1*) promoter of maize (Christensen *et al.,* 1992; Cornejo *et al.*, 1993). It should be noted that endogenously the *Ubi1* promoter is not considered constitutive. It activates expression of the respective gene especially in response to some stresses (Christensen *et al.,* 1992). This renders this promoter very efficient during the transformation procedure where the transformed tissue is under stress. Thus, expression of the selectable gene is at the period when the latter should be expressed to assure the selection of transformed cells.

The *Ubi1* promoter activity is also observed to be environmentally responsive in maize (Christensen *et al.*, 1992) and it responds to a range of stresses in transgenic sugarcane (Hansom *et al.*, 2001). Under field conditions, this stress responsiveness may account for a degree of variation in transgene activity levels between adjacent stalks of the same transformed plant line (Hansom *et al.*, 2001). The variable *Ubi1* promoter activity can be a disadvantage for the practical application of many potentially useful new genes in sugarcane since stability or predictability of transgene expression over time would be a vital consideration. In addition, specific patterns of transgene product accumulation are required and necessary to fulfill the intended agronomic function of each new gene in the

sugarcane plant (Christopher and Campbell, 2001). Thus, one area of improvement would be to explore sequence alterations within the *Ubi1* promoter for reduced inducibility or more constitutive properties (Hansom *et al.*, 2001).

2.3.1 The regulatory role of introns in gene expression in monocots

The coding region of many eukaryotic genes is interrupted by non-coding intervening sequences known as introns (Freeling and Bennet, 1985). They are removed from the nascent mRNA in a multistep process collectively called splicing (Mogen *et al.,* 1992). As reported previously, the presence of introns in the 5' untranslated region of the gene is shown to be an important component for expression in plants.

The 900 bp fragment of the maize *Ubi1* promoter at the 5' end of the gene contained in a genetic construct, drove expression of the chloramphenicol acetyl transferase (*Cat*) gene ten fold higher than the cauliflower *35S-Cat* in maize protoplasts. However, it was observed that the maize *Ubi1* promoter did not drive transgene expression in the protoplasts of tobacco, a dicot (Christensen *et al.*, 1992).). Although transcriptional control elements have been localised to introns (Vasil *et al.*, 1992), the requirement of the intron was not a function of the promoter or transcription initiation site. It was rather a function of the structure and/or processing of the transcribed sequences in producing mature cytoplasmic mRNA through splicing mechanisms (Freeling and Bennet, 1985).

Furthermore, intron-exon splice junctions are different between monocots and dicots (Simpson *et al.*, 1996). Monocot junctions are predominantly pyrimidine rich bases in the 3' splice junction, whilst the dicot 3' splice junctions are higher in purine (Hanler and Schuler, 1988). The first intron of a wheat ribulose-1, 5-bisphosphate carboxylase-oxygenase (*rbcS*) gene was found to be inefficiently spliced in transgenic tobacco, a dicot (Keith and Chua, 1986). Hence, molecular interactions at these specified junctions could be the major determinant in

processing mature mRNA transcripts (Hanler and Schuler, 1988; Keith and Chua, 1986).

Alternatively, transcriptional control elements present in introns can make promoters such as maize *Ubi1* and maize *Adh1,* monocot specific (Christensen *et al.*, 1992). They provide binding sites for transcription factors that are lacking in dicot systems, and consequently affect gene expression. An example is the barley ABA responsive *Hva22* gene that encodes a potential regulatory protein. One of the elements essential for ABA induction is present within the intron (Shen and Ho, 1995). In addition, homologies to known transcription factor binding boxes such as an *opaque*-2 binding motif (GATGAYRTGR) (Lohmer *et al.,*1991) and an E-box (CANNTG) can be found at positions +404 and +452 within the maize ubiquitin intron sequence, respectively (Salgueiro *et al.,* 2000). Also identified are promoter-like features such as a TATA box like sequence at position +924 and a canonical CAT box at position +390 (CAAT), much further upstream from the TATA box than usually found in plant promoter sequences (Salgueiro *et al.,* 2000). In contrast, the maize *Ubi1* promoter does not have a CAT box (Christensen *et al.,* 1992). Thus, these elements may play a role in the ability of these introns to drive gene expression to a low extent. They may also have a regulatory role when the introns are part of a transcriptional cassette, in the absence of known regulatory promoter elements (Salgueiro *et al.,* 2000). This further confirms the regulatory role of the 5' and 3' upstream regions present in the leader sequences of genes (Christensen *et al.*, 1992).

2.4 The Importance of Specific rather than Constitutive Promoters

To allow the manipulation of key regulatory components of important metabolic pathways in the sugarcane, it is desirable that the expression of transgenes in genetically modified crops is restricted to the tissues requiring the encoded activity (Christopher and Campbell, 2001). Tissue and developmental stage specific promoters from sugarcane would also relieve the plant from the strain of

expressing the transgene throughout plant growth and development (Christopher and Campbell, 2001). The lack of temporal and spatial regulation of a constitutive promoter such as the CaMV *35S* may be suitable for proof of concept experiments, but it has a number of potential drawbacks for use in genetically modified crops (Gittins *et al.,* 2000). For instance, the presence of multiple transgenes driven by the same constitutive promoter in a single plant (e.g for resistance gene-pyramiding) may result in homology dependent gene silencing, particularly where the promoter is endogenously highly active (Matzke *et al.,* 1994). There is also the potential risk that expression of viral capsid proteins via constitutive promoter in transgenic plants, may increase the risk of transcapsidation or viral recombination to generate new strains of phytopathogens (Gittins *et al.*, 2000).

Since the compartmentation of key enzymes involved in starch and sucrose synthesis has not been unequivocally determined in sugarcane, targeting genetic manipulation through specific promoters may provide a clearer understanding of the sucrose pathway. For example, by targeting gene expression in the vacuoles of the stem parenchyma, enzyme activity of acid invertase, the principal enzyme involved in the hydrolysis of the sucrose, can be modified through co-suppression mechanisms (Albert *et al.*, 1996). In the apoplast tissues of transgenic tobacco plants, expression of yeast invertase was localised to manipulate carbon partitioning (Albert *et al.*, 1996). Results showed the direct relationship of sucrose for growth and development in sink tissues to their transport via the phloem. By regulating the principal rate or co-limiting steps in the entire sucrose accumulation process, the crop can be fully exploited for increased sucrose content in the stem.

Specific promoters have been characterised in other plants. For example, sucrose synthase enzyme from maize, one of the key enzymes in sucrose metabolism, has been identified (Yang and Russel, 1990). Its promoter activity is phloem cell specific in transgenic tobacco plants. The ribulose-1,5-bisphosphate carboxylase (Rubisco) small subunit (SSU) gene promoters are active primarily in the green photosynthetic tissues of apples (Gittins *et al.*, 2000). Furthermore, the grape

alcohol dehydrogenase promoter is expressed during ripening of the grape berry and hence characterised as a fruit-ripening promoter (Sarni-Manchado *et al.*, 1997).

2.4.1 Characterisation of regulatory domains conferring specific promoter qualities

2.4.1.1 Control in tobacco plants

In addition to the use of the CaMV *35S* in plant gene expression vectors, this promoter offers a convenient system to dissect promoter organisation and to identify functional cis-acting elements responsible for tissue and developmental-stage specific expression in higher plants (Benfey and Chua, 1990). Three types of approaches have been carried out to identify control motifs within its promoter region. Promoter elements are dissected through: (i) loss of 5', 3'ends and internal deletions, (ii) gain of functional analysis with different promoter regions, and (iii) site specific mutation of known transcription factor binding sites (Benfey and Chua, 1990).

A detailed analysis of the *35S* promoter has been carried out with transgenic tobacco plants by Fang *et al.* (1989). Results from a combination of 5' and 3' deletions suggested the following model of functional architecture for the *35S* promoter: (1) the region from the –343 to –46 upstream contains at least three domains of functional importance for leaf expression. The domain between –343 and –208 is responsible for about 50% of the promoter activity. Deletions of sequences between –208 and –90 further decrease the remaining activity about two-fold. The third domain, from –90 to –46 is required for the synergistic activation observed with upstream sequences (-343 to –107); (2) the region from –107 to –78 can also interact with upstream sequences to drive high levels of expression; (3) promoters truncated at –90 or –105 have little activity above background (Benfey and Chua, 1990). These results show that for maximal levels of expression in tobacco leaves, multiple sequence elements between –343 and –46 are required

and synergistic interactions between elements are critical as was demonstrated by the inter-dependence of the sequences. From the above deletion analysis, it became apparent that the *35S* promoter is a composite of *cis* acting elements that have distinct functional properties.

To establish a quantitative model for the tissue specificity of sub-domains within this promoter, various upstream regions were tested for activity in transgenic tobacco and petunia (Benfey and Chua, 1989; 1990). The coding region of the bacterial β-glucuronidase (GUS) was used as a reporter gene in order to detect gene expression by histochemistry. Using this reporter gene, the –90 truncated *35S* promoter was found to be active in roots of transgenic tobacco while the upstream sequences (-343 to –90) are preferentially expressed in leaves (Benfey and Chua, 1989).

The region between –90 to –72 is critical for the observed activity in roots. The sequences between –343 and –90 were further divided into five sub-domains and their activities were studied individually in transgenic plants (Benfey and Chua, 1990). The results showed that each sub-domain is capable of activating gene expression in a distinctive manner but each of them exhibit low activity when placed upstream of the –46 truncated promoter. Hence, the interdependence of different *cis* acting elements to drive cell-type specific expression illustrates the combinationial nature of promoter elements. The results also indicate that the –90 to –46 region of the *35S* promoter must contain elements that can interact with various types of upstream elements.

2.4.1.2 Tuber specificity and sucrose induction of patatin expression

Other regulatory elements from tissue specific genes include the patatin promoter of potato (Wenzler *et al.*, 1989). Patatin is a lipolytic acylhydrolase (Andrews *et al.*, 1988), which may release fatty acids from membranes as part of a defense

response. Two classes of genes (I, II) encode for the expression of this enzyme. Class I is expressed at high levels in tubers, while Class II is expressed at low levels. High expression of patatin could be regulated by tuber specific factors or by the high concentration of sucrose. Analysis of its promoter organisation through promoter studies in transgenic potato, revealed a high expression regulated by a minimal promoter extending 344 bp distal to the transcription start site (Wenzler *et al.*, 1989).

Within this region, a tuber specific enhancer was defined (Jefferson *et al.*, 1990). Two sequence motifs termed A and B repeats were also identified. Increased expression was observed with longer promoters with additional A and B repeats. Furthermore, different regions were found to be sucrose responsive or tuber specific. A deletion of B repeats between –224 and –207 lowered tuber specificity from 167 fold to 10 fold (Wenzler *et al.*, 1989). It also lowered sucrose inducibility by half. However, there is evidence that sucrose itself is not directly involved in signalling the activation of patatin transcription via transcription factors SURF and BBF (Wenzler *et al.*, 1989). There is a possibility that the flux of metabolites from sucrose to starch is detected and that signals patatin transcription.

2.4.1.3 Abscisic acid induced *cis* acting elements

During the growth and development of plants, abscisic acid (ABA) modulates gene expression, particularly during seed formation and in response to environmental stress involving loss of water. An approach to study the molecular mechanisms of ABA induction was to identify *cis* acting elements necessary and sufficient for ABA response as well as the transcription factors interacting with these elements. Examples of ABA induced genes include Em wheat protein that accumulate in the last stages of grain development (Marcotte *et al.*, 1989).

From studies based on transient GUS assays in rice protoplasts, six copies of the sequence motif GTACGTGGCGC were identified within the Em wheat promoter

that conferred six-fold ABA induction to a CaMV *35S* minimal promoter. This sequence and its homologues have been designated as ABRE (ABA response elements)(Shen and Ho, 1995) and these are very similar to the G-box with a common ACGT core, a regulatory element responsive to physiological factors such as light (Giuliano *et al.,* 1988) and auxin (Liu *et al.,* 1994).

Recent studies on the barley ABA responsive *Hva22* gene, analysed by both loss and gain of function studies, revealed that G box sequences are necessary but not sufficient for ABA response (Shen and Ho, 1995). Instead, an ABA response complex consisting of a G box, namely ABRE (GCCACGTACA) and a novel coupling element CE1 (TGCCACCGG) was found to be sufficient for high level ABA induction (Pla *et al.,* 1993). In an ABA responsive gene more than one ABA response complex can be present. The two complexes of the *Hva22* promoter contribute to the expression level of the gene in response to ABA. The flanking sequence around the ACGT core may also participate in determining the signal response specificity by differentiating the interactions with various ABRE or G box binding proteins. The presence of more than one signal response complex is not unique to the ABA responsive *Hva22* gene and is also found in the light responsive elements of chalcone synthase of the *Chs*-genes (Schultze-Lefert *et al.,* 1989), and in the auxin inducible GH3 gene (Liu *et al.,* 1994).

2.5 Differentially Expressed Genes as Research Tools for Specific Promoter Isolation

Since the genes that are part of the multigene families are believed to have evolved mechanisms that prevent them from interaction during plant development (Mittelsten Scheid *et al.*, 1996), an attractive alternative would be to isolate promoters of interest from the plant itself. Since the mature sugarcane culm stores high concentrations of sucrose and one of the main goals of sugarcane biotechnology is to obtain an understanding of sucrose metabolism, isolation of

promoters from that region of the plant would be a better option rather than using foreign promoters for sugarcane genetic modification.

The collection of genes that are expressed in a tissue at that specific stage are responsible for both morphological and phenotypic characteristics and they are indicative of cellular processes to environmental stimuli and perturbations (Jacobsen *et al.,* 1992). Thus, differentially expressed genes in sugarcane can be used to identify tissue or stage specific promoters, which can be used for genetic manipulation of the crop.

By comparing the transcript abundance of individual mRNA's present in samples originating from different genotypes, developmental stages or growth conditions, genes can be identified that are differentially expressed. Isolation of such genes could allow the associated promoters to be sequenced and compared. As reported previously, common *cis* regulatory elements may be localised and their presence may be correlated with specific features of the expression profiles of the corresponding genes (Nover, 1994). The latter may represent specific metabolic and morphogenetic functions.

Numerous genes and in some cases their associated promoters, which exhibit a wide range of tissue and/or developmental expression patterns, have been characterised. Examples consist of genes that are specifically expressed in pollen (Custers *et al.*, 1997), embryos, cotyledons, endosperm, phloem (Clark *et al.*, 1997), roots, fruit (Coupe and Deikman, 1997), as well as symbiosis related genes in myccorhizas (Martin-Laurent *et al.*, 1997). A deepened insight into the activity of such genes provided information on key regulatory mechanisms e.g. cell differentiation, osmoregulation (Smart *et al.*, 1998) and sucrose metabolism (Heineke *et al.*, 1992). Due to the limited knowledge of the sugarcane genome, a collection of such data remains unavailable. However, there is reason to believe that genes are differentially expressed in the sugarcane plant. Studies based on

the histochemical staining and sugar analysis of the sugarcane have revealed the following:

i.) An increase in the number of vascular bundles but a decrease in size is observed from the core to the peripheral tissues within a sugarcane internode (Jacobson *et al.,* 1992).

ii.) An increase in the pattern of lignification and suberisation was observed down the mature culm, which also paralleled an increase in sucrose storage (Jacobson *et al.,* 1992).

Hence, differences in gene expression are responsible for the variation in the structure, morphology and metabolic activity in the sugarcane culm. This finding formulates the hypothesis of this study that genes are differentially expressed in the sugarcane culm at different developmental stages. Since messenger RNA (mRNA) is only an intermediate on the way to production of protein products, the question is why measure mRNA at all?

Protein based approaches are generally more difficult, less sensitive and gives a lower throughput than RNA based ones. But more importantly, mRNA levels are highly informative about cell state and the activity of genes, and for most genes, changes in mRNA levels correlate with changes in protein abundance (Lockhart and Winzeler, 2000). To evaluate level of transcripts, abundance of mRNA, gene expression and changes in gene expression, many techniques have been devised. Northern hybridisation analysis, polymerase chain reaction after reverse transcription of RNA (RT-PCR), nuclease protection, cDNA sequencing, clone hybridisation, differential display, subtractive hybridisation, cDNA fragment fingerprinting, serial analysis of gene expression and array technology have all been put to good use to measure expression levels of specific genes and compare their expression profiles (Wang *et al.*, 1999). The molecular principles of a few of the methods are described in the next section, citing their significant application in characterising differentially expressed genes.

2.5.1 Differential screening of cDNA libraries

Differential screening of cDNA libraries is one of the conventional methods used for the identification of differentially expressed genes in an eukaryote. This technique involves the differential screening of plasmid or phage cDNA libraries. However in practice, this method is only effective if the mRNA of interest consists of more than 0.05% of the total mRNA in one cell and less than 0.01% in the other (Sambrook *et al.*, 1989).

The sensitivity in targeting rare mRNA transcripts is increased if the concentration of the specific sequence is enriched via subtractive hybridisation (SH) (Aguan *et al.*, 1991). The mRNA extracted from one cell type is used as a template to synthesise radiolabelled cDNA which is then hybridised exhaustively with an approximately 20 fold excess of mRNA extracted from a second cell type in which the gene of interest is not expressed (Aguan *et al.*, 1991).

Complementary nucleic acids form DNA:RNA hybrids, which are then separated from the single stranded cDNA by chromatography on hydroxylapatite columns or streptavidin biotin interactions. The subtracted cDNA fragments can then be used for screening or even for construction of subtracted cDNA libraries. Rare mRNA's with prevalences ranging from 1/500 to 1/200 000, with the median being approximately 1/20 000, can be detected with this procedure (Hodge *et al.*, 1992).

Application of this technology led to the characterisation of phytochrome genes specific to the flower development in tobacco and male flower from maize (Wan *et al.*, 1996).

2.5.2 Array technology

The array technology used to measure gene expression involves the deposition of denatured double-stranded DNA fragments on a matrix. The immobilised

molecules on the array surface are then hybridised to different mRNA populations derived from different sources in individual experiments. Expression profiles of the arranged DNA is then compared to identify genes that are constitutively expressed or differentially expressed from the two mRNA samples. Single DNA probes are then prepared by a reverse transcriptase of poly $(A)^+$ RNA in the presence of ^{33}P – dATP to measure transcript concentrations by hybridisation to the arrays (Chen *et al.*, 1998).

The availability of sophisticated image processing software allows the identification and quantification of hybridised signals as well as subtraction of local background values. The use of radioactive isotopes can be replaced by using a non-radioactive single or dual colour detection system in which gene expression is represented by colour intensities generated by colorimetric reactions (Chen *et al.*, 1998). The array technology comprises of the macro-array and the micro-array. The macro-array allows more than 6000 elements at indexed locations on nylon membranes using high speed arraying machines, although it may be done manually (Chen *et al.*, 1998).

In contrast, the micro-array can contain 20 000 cDNA targets on microscope slides deposited by robotic printing. However, the source of cDNAs for both macro-array and micro-array could be known sequences (EST's) or anonymous cDNA clones derived from a cDNA library. The usefulness of cDNA macroarrays for monitoring gene expression in higher plants has been demonstrated by Desprez *et al.* (1998), who used a collection of 432 partially sequenced cDNAs arrayed on a nylon support.

The filters were employed in hybridisation studies with ^{32}P labelled complex probes prepared from mRNA of light- and dark- grown *Arabidopsis thaliana* seedlings. Only a small number of mRNAs, corresponding to 1% of the cDNAs tested, were found to be represented at levels higher than 0.5% of the total mRNA level. Furthermore, comparison of light- and dark- grown seedlings showed significant

differences in the expression level for approximately 15% of the genes investigated (Desprez *et al.*, 1998).

In addition, accurate quantification technology with the microarray technology revealed the occurrence of many genes for which differential patterns of expression had not been observed previously under similar conditions. Examples of identified differentially expressed genes include the chlorophyll a/b binding proteins, plastcyanin and ribulose-1,5-bisphosphate carboxylase (Kuhn, 2001). Furthermore, an example of microarray analysis is reported in the identification of differentially expressed genes involved in different organs and at different developmental stages of the *Arabidopsis thaliana* plant (Schena *et al.*, 1995). 1443 *Arabidopsis thaliana* genes were assayed and more than 80% of the DNA elements on the arrays generated detectable fluorescence signals (Schena *et al.*, 1995).

However 1,4 to 5 % represented highly expressed genes with an abundance of more than 100 – 500 transcripts per cell. The most differentially expressed sequences were then known to be involved in photosynthesis, showing up and down regulation patterns of gene expression from leaves, roots and flowers. Novel sequences were also identified showing no homology to identified sequences in the established database (Schena *et al.*, 1995).

2.5.3 Oligonucleotide array hybridisation analysis

The basic principle of using oligonucleotide microarrays was first proposed in the late 1980's, when several groups independently developed the concept of sequencing by hybridisation, i.e. determining a DNA sequence by hybridisation to a comprehensive set of oligonucleotides, such as all the possible 65536 octamers (Chetverin and Kramer, 1994). The initial aim was the establishment of a faster, less costly and more effective sequencing approach. When it gradually became apparent that microarrays could also be used for large-scale comparative

measurements of mRNA levels of sets of genes in different cell types or under different conditions, the interest shifted to transcriptional profiling, which is now the most prominent application (Chetverin and Kramer, 1994).

By automatically printing presynthesised oligonucleotides onto microscopic slides, oligonucleotide microarrays can be produced in the same way as cDNA microarrays at densities of up to 30 000 elements.cm^{-2} (Yershov *et al.*, 1996). A different way of producing oligonucleotide microarrays is semiconductor photolithography in combination with modified phosporamidite-based oligonucleotide synthesis chemistry (Foder *et al.*, 1991). This technique generates high-density microarrays using comparatively few synthesis steps, since the number of probes increases exponentially with a linear increase in the number of synthesis cycles. For example, a complete set of octanucleotides consisting of 65 536 probes can be produced in 32 chemical steps in 8 hours (Pease *et al.*, 1994). The basic steps of probe labelling, hybridisation, signal detection and data processing are the same regardless of array type. Oligonucleotide microarrays have been used to measure expression levels of genes in mice (Lockhart *et al.*, 1996), yeast (Wodicka *et al.*, 1997) and bacteria (DeSaizieu *et al.*, 1998).

2.5.4 Serial analysis of gene expression

Serial analysis of gene expression (SAGE) is another large scale method that allows quantitative and simultaneous analysis of many transcripts (Velculescu *et al.*, 1995). From each species present in a cDNA population, a unique sequence tag (approx. 9bp) is extracted by a series of enzymatic restriction and ligation steps. The cleavage reactions are catalysed by a pair of restriction enzymes, a tetracutter termed the anchoring enzyme and a type II S restriction endonuclease, which is termed the tagging enzyme (Velculescu *et al.*, 1995).

The tags are linked together, concatenates are inserted into a plasmid vector, cloned and sequenced. In the sequence, tags are easily distinguished because

they are separated by the recognition sequence of the anchoring enzyme, which establishes the register and boundaries of each tag. Hence concatenation of short sequence tags allows the efficient analysis of transcripts in a serial manner by the sequencing of multiple tags within a single clone. A transcription profile is obtained by determining the total frequency of each tag in a SAGE library, which directly reflects the abundance of the corresponding mRNA in the tissue (Velculescu *et al.*, 1995).

2.5.5 The cDNA-AFLP technique

This technique involves the RNA fingerprinting technique and it evolved from the amplified fragment length polymorphism (AFLP), as described by Vos *et al.* (1995), for the fingerprinting of genomic DNA. The cDNA-AFLP involves three steps:

i.) restriction of cDNA and ligation of oligonucleotide adapters,
ii.) selective amplification of restriction fragments using PCR primers bearing selective nucleotides at the 3' end, and
iii.) gel analysis of the amplified fragments.

Generally, restriction of plant cDNA with a combination of two restriction enzymes, a tetracutter and a hexacutter provide a significant cDNA population to be cleaved and represented as a discrete banding pattern on a sequencing gel. A subset of the cDNA fragments generated is then amplified using PCR primers with 2 selective bases for each primer, giving a total of 256 possible primer combinations. This gives a reasonable banding pattern of cDNA fragments that is visualised by autoradiography. The cDNA-AFLP products visible on a polyacrylamide sequencing gel range between 100 to 1000 bp. In this range, an average of 40 bands can be observed for each primer combination, corresponding to a total of approximately 10 000 bands. This method is a quantitative high throughput expression profiling method specifically designed to measure the concentrations of known transcripts in numerous different tissues (Vos *et al.*, 1995).

2.5.6 Differential display reverse transcription PCR (DDRT – PCR)

DDRT – PCR, the Differential Display Reverse Transcription PCR was developed by Liang and Pardee in 1992. The method was intended to provide an effective tool to detect individual mRNA species that are differentially expressed in different eukaryotic cell types and then to permit recovery and cloning of the corresponding cDNAs.

The principle of the DDRT – PCR technique is to target specific subpopulations of mRNA using subsets of 12 anchored oligo (dT) primers. These primers bind to different fractions of the total mRNA, which are then reverse transcribed. The resulting cDNA fragments are then amplified using the same-anchored oligo (dT) and a short arbitrary primer. Labelled dATP is introduced in the reaction and the labelled products are separated on a DNA sequencing gel prior to visualisation by autoradiography. Many refinements of the original differential display technology have been described (Donohue *et al.*, 1995). Instead of decameric upstream primers, elongated arbitrary primers have been used. This opens the possibility of including common sequence motifs to target particular classes or families of genes (Donohue *et al.*, 1995; Gonsky *et al.*, 1997). Methods employed for the separation of DDRT – PCR products include denaturing and non-denaturing polyamide gel electrophoresis, capillary electrophoresis and even electrophoresis on standard agarose gels (George *et al.*, 1997). For labeling of DDRT-PCR products, [^{33}P]dATP is generally used. However, alternatives to radioactive labeling are fluorescence detection (Bauer *et al.*, 1993), silver staining (Gottschlich *et al.*, 1997) and chemilumiscence detection (An *et al.*, 1996).

2.5.7 RFLP coupled differential display

Many genes share a similar modular structure or certain domains in their DNA sequence and protein product. These features are known as family specific domains, (FSDs). These FSDs classify genes into groups of gene families and are

well characterised for the chlorophyll a/b binding proteins as well as for many transcription factors (Kuhn, 2001).

Restriction fragment length polymorphism coupled domain directed differential display (RC4D) is a method specifically designed to analyse expression of multigene families at different developmental stages, in diverse tissues or in different organisms (Fischer *et al.*, 1995). RC4D combines cDNA – AFLP technology with a gene family specific version of DDRT – PCR.

In RC4D, longer primers directed against FSD instead of decameric primers are used, allowing cDNA's belonging to the same gene family to be selectively amplified (Kuhn *et al.*, 2001). As the amplification products are relatively uniform in length, restriction fragment length polymorphism (RFLP), is introduced by digestion with a frequently cutting restriction enzyme. The amplicon size from 1 kb to 100 bp is thus reduced, optimal for separation on acrylamide gels. Hence, family members can easily be distinguished by size (Kuhn, 2001). Application of this technology include characterisation of MADS lux genes, which are differentially expressed in male and female inflorescence of maize (Fischer *et al.*, 1995).

2.6 Reproducibility, Accuracy and Throughput Analysis of Display Techniques

cDNA array hybridisation, oligonucleotide array hybridisation, AFLP and serial analysis gene expression (SAGE) represent quantitative high throughput display techniques allowing the intracellular concentrations of mRNA transcripts to be determined simultaneously in a single hybridisation experiment. The macro-array technology is relatively cheap with filters that can be hybridised up to ten times (Desprez *et al.*, 1998). However, micro-arrays are fifty times more sensitive compared to the macro-arrays, allowing two mRNA populations to be analysed at one time. Since no sequence characterisation is required for genes of an unknown

genome, anonymous cDNA clones derived from a cDNA library can be used for array preparation.

However, a very large number of clones on the arrays is required to ensure that a low abundant mRNA is detected. It has been reported that more than 20 standard size cDNA microarrays are required in order to have a 99% probability of at least one array element hybridising to a rare mRNA. This makes the array technology an incompetent means for the identification of rare mRNA transcripts corresponding to unknown genes, despite its high potential for automation (Kuhn, 2001)

The power of SAGE is comparable to microarray technology. However, low abundant transcripts and small differences in transcript concentration can be detected using this technique. Two major advantages of the SAGE over the cDNA expression arrays are that firstly, uncloned cDNA can be analysed and secondly, no other special device is required other than a sequencer. Furthermore, a representative SAGE library much smaller than a cDNA library, is required to build an equivalent microarray. This is because the SAGE clones may contain 20 to 50 tags (Datson *et al.*, 1999).

On the other hand, the cDNA-AFLP, DDRT-PCR and differential screening techniques are known for their capacity to analyse transcription profiles on a qualitative basis. The minute amount of RNA needed in the preamplification step of the cDNA-AFLP technique is advantageous due to the easily degradable aspect of the single stranded RNA isolated from tissue samples. Because stringent hybridisation conditions are used in the amplification reactions, mismatched priming events are observed in cases where transcripts are extremely high (Bachem *et al.*, 1996). This results in the cDNA-AFLP banding patterns being highly reproducible and almost free of false positives. A straightforward verification mechanism of band identity and homogeneity can be established via this procedure. When using AFLP primers with additional base extensions, it is possible to selectively amplify one cDNA fragment, thereby eliminating virtually all

other fragments from being co-amplified. Transcript concentrations can be measured quite accurately by combining cDNA-AFLP with high-resolution quantitative separation techniques (Bachem *et al.*, 1996). This is because band intensity is observed to be a direct function of template concentration. However, a significant disadvantage of this technique is the requirement of appropriate restriction sites on the cDNA molecule. In order to visualise every cDNA in a plant cell, different enzymes are required. Furthermore, the technical simplicity of the DDRT-PCR has led to the identification of several differentially expressed genes. A potential drawback to this technology is the considerable number of false positives generate in experiments (Gonsky *et al.*, 1997).

Lastly, with the many steps required to construct a representative cDNA library and the necessity to screen thousands or millions of clones, differential screening is considered to be extremely laborious and inefficient to identify rare mRNAs (Kuhn, 2001). However, genes that are highly expressed in a cell under a set of conditions but weakly expressed under others, can be easily detected with a small number of clones.

Thus, as reported, all differential display techniques provide information about steady state mRNA concentrations. However, with the wide range of display techniques available, the choice of method to investigate the pattern of gene expression is mostly dictated by cost, type of problem to be solved and the amount of sequence information available for the species under study.

Since not much information is available on specific genes in the sugarcane culm, the approach of this study would be to target random genes expressed abundantly in specific tissues or stages of the mature culm. No prior sequence characterisation of the genes of interest are required. Hence, the differential screening of a late stage cDNA library combined with the macroarray technology has been chosen for this investigation and may be a valuable way of identifying differentially expressed

genes in the mature culm. Strong stage and tissue specific promoters could then be isolated using the differentially expressed genes as research tools.

CHAPTER 3
Materials and Methods

3.1 Plant Material and Sample Preparation

Randomly selected field-grown sugarcane plants (variety N19) were harvested over the summer season from the Welgevallen farm in Stellenbosch. Plants with approximately 15 internodes were selected. The internode attached to the leaf with the uppermost visible dewlap was defined as internode number 1, according to the system of Kuijper (Van Dillewijn, 1952). In the laboratory, internodes 2, 7 and 10 were excised from the sugarcane stalks and the rind removed. The 5mm peripheral tissues under the rind and the 5mm inner core tissues were cut up from each internode and frozen in liquid nitrogen. Tissue samples were stored at -80^0 C.

3.2 Isolation of RNA

RNA is degraded by ubiquitous ribonucleases (RNases) (Croy *et al.*,1993). The following precautions were taken when working with RNA. Latex gloves were worn at all times when handling RNA samples and solutions that were used for RNA extraction experiments. Glassware, spatulas eppendorf tubes and micropipette tips were baked at 180^0C for 6 h, for the inactivation of ribonucleases, prior to usage. Glass centrifuge tubes, solutions, mortar and pestle were treated by overnight incubation at 37^0C with 0.1% (v/v) diethylpyrocarbonate (DEPC) treated water. All RNA solutions were prepared using DEPC treated water except for Tris buffer. Solutions and glassware were then autoclaved to destroy residual DEPC. Stocks of chemicals and disposable plasticware for RNA work were kept separate.

3.2.1 RNA extraction and poly $[A]^+$ RNA selection

Total cellular RNA was extracted using two methods:
Method A: Perchlorate method with a few modifications (Davies and Robinson, 1996).

Method B : A phenol based method

3.2.1.1 Perchlorate Method

Frozen plant tissues were ground into a fine powder using a mortar and pestle that had been pre-cooled with liquid nitrogen. Four grams of homogenised tissue were added to a 50ml Corning tube on ice. To each sample, 20ml of RNA extraction buffer (5 M sodium perchlorate; 300 mM Tris-HCl, pH 8.3; 1% (v/v) sodium dodecyl sulphate (SDS); 10% (w/v) polyvinyl poly-pyrrolidone (PVPP); 2% (v/v) polyethylene glycol 6000 (PEG-6000); 1% (v/v) β-mercaptoethanol) was added. The tissue powder was carefully dispersed in the solution and the samples were then placed on a rotary shaker for 30 min and vortexed every 10 min, to ensure mixing. The slurry was then transferred to glasswool filters and centrifuged at 792*g* for 20 min at 4^0C. Glasswool filters were made by lightly stuffing the bottom of a 20ml syringe with sialinized glasswool (Ausubel, 1996) until it was around 1.5 cm thick. This syringe was then placed inside a 50ml Corning tube and used as a filter.

To the eluate, 2.5 volumes of 100% (v/v) ethanol was added and the mixture was left to incubate overnight at -20^0 C. Samples were then centrifuged at 8000*g* for 20 min at 4^0C. The supernatant was discarded and the pellet rinsed with ice-cold 70% (v/v) ethanol. The pellet was suspended in 2ml TE buffer (10 mM Tris-HCl, pH 7.6; 1 mM ethylene diamine tetra-acetic acid disodium salt, EDTA) containing 0.2% (v/v) β-mercaptoethanol. The solution was divided into 3 microcentrifuge tubes prior to phenol:chloroform:isoamylalcohol (25:24:1) extraction. This was followed by a chloroform : isoamylalcohol (24 : 1) extraction. Samples were then centrifuged at 8000*g* for 10 min at 4^0 C. This step was repeated until the interphase between the organic and aqueous phases was clear of proteins. The aqueous layer was then removed, added to 2.5 volumes of 100% (v/v)

ethanol and 0.1 volumes of 3 M sodium acetate, pH 5.2, and incubated at –20^{0}C for 2 h. The samples were then centrifuged at 18000*g* for 20 min at 4^{0}C, the supernatant discarded and the pellet rinsed with 70% (v/v) ethanol.

The pellet was then resuspended in 200µl DEPC treated water by incubating at 65^{0}C for 10 min. To the suspension, 0.3 volumes of 8 M lithium chloride was added and the RNA was allowed to precipitate for 12 h at 4^{0}C. After centrifugation at 18000*g* for 30 min at 4^{0}C, the supernatant was discarded, the pellet was dried at room temperature and resuspended in 250µl DEPC treated sterile water by incubating at 65^{0}C for 15 min. To remove particulate matter, the samples were centrifuged at 3200*g* for 5 min at 4^{0}C and the supernatant of all the tubes was combined to form one sample. RNA solutions were then stored at –80^{0}C.

3.2.1.2 Phenol based Method

For this method, 2.8 grams of frozen ground tissue was added to a 50ml Corning tube containing 16ml of RNA extraction buffer (1% (w/v) sodium dodecyl sulphate; 100 mM Tris-HCl, pH 7.5; 100 mM sodium chloride; 1 mM EDTA) and 8ml of phenol, which was being held at 70^{0}C in a water bath. After 10 min, 5.6ml of chloroform was added to each sample and the slurry was centrifuged at 8000*g* for 15 min at 4^{0}C. The aqueous phase was then transferred to a clean centrifuge tube and 0.1 volumes of 3 M sodium acetate was added. This was followed by the addition of 3 volumes of 100% (v/v) ice cold ethanol for RNA precipitation at –20^{0}C for 2 h. After centrifugation at 8000*g* for 30 min at 4^{0}C, the supernatant was discarded and the pellet rinsed with ice cold 70% (v/v) ethanol. The RNA was resuspended in 10ml of DEPC treated water and adjusted to 2 M lithium chloride (final concentration) for RNA precipitation at 4^{0}C overnight. The RNA was then recovered by centrifugation at 1000*g* for 30 min at 4^{0}C and the

supernatant discarded. The RNA pellet was then rinsed with 70% (v/v) ethanol, air dried and resuspended in 250µl DEPC treated water.

To determine yield and integrity of extracted RNA, all RNA samples were diluted 50 times and quantified by ultraviolet (UV) spectrophotometry. A standard assay for assessing RNA purity was performed by calculating the ratio of sample absorbancy at 260 nm and 280 nm. The A_{260} values were then used to determine the RNA concentration according to the Lambert-Beer Law equation: A = ECl, where A is the sample absorbancy at 260 nm wavelength, C is the concentration of RNA ($\mu g.ml^{-1}$), l is the pathlength of the spectrophotometer cuvette (1 cm) and E is the single-stranded RNA specific adsorption coefficient value, corresponding to 0.025 ($\mu g.ml^{-1}$) cm^{-1} (Ausubel, 1996).

The RNA quality was assessed by gel electrophoresis of samples through 1.5% (w/v) agarose gels. Samples were prepared by adding 50% (v/v) formamide, 6% (v/v) formaldehyde and 1x MOPS buffer (200 mM 3-[N-morpholino]propanesulphonic acid (MOPS); 50 mM NaOAC; 5 mM EDTA) and heating to 65^0C for 10 min before loading. For visualisation purposes, one sample was loaded separately together with the 0.24 – 9.5 kb RNA Ladder (GibcoBRL) and these were treated the same as the RNA samples. After separation in the gel, the ethidium bromide ($0.1\mu g.ml^{-1}$) stained RNA bands were visualised using the Geldoc system (Alpha Imager). After checking for its integrity, a range of 120–130 µg of total RNA from each tissue was used to extract mRNA, using the Poly A Tract mRNA isolation kit, according to the manufacturer's instructions (Promega).

3.3 Construction of Internode 7 and Internode 10 cDNA Library

3.3.1 cDNA synthesis

First and second strand synthesis were performed according to the method and conditions described in the protocols and applications guide of the

Universal Riboclone® cDNA synthesis system (Promega). Approximately 1μg of poly $[A]^+$ RNA was used in first strand synthesis reaction catalysed by the AMV (Avian Myeloblastosis Virus) reverse transcriptase enzyme with synthetic oligonucleotide, oligo $(dT)_{15,}$ as a primer. Components of the second strand cDNA synthesis were added directly to the same tube following first strand cDNA synthesis.

After heat inactivation at 70^0C for 10 min, the second strand cDNA reaction was completed by the addition of T_4 DNA polymerase ($2U.\mu g^{-1}$ mRNA) and incubated for 10 min at 37^0C. The double-stranded (ds) cDNA product was then phenol:chloroform extracted and purified through a QIAquick-spin PCR purification column (Qiagen GambH, Hilden, Germany) according to the manufacturer's instructions. cDNA was ethanol precipitated prior to ligation to amplification adaptors.

3.3.2 Ligation to amplification adaptors

Ligation to amplification adaptors was performed according to the protocol described in the Promega Protocols and Applications Guide. The cDNA was blunt end ligated to an annealed amplification adaptor set (Jepson *et al*, 1991). This adaptor set consisted of the following two oligonucleotides:

Oligonucleotide 1 (29-mer):

5' - ATGCTTAGGAATTCCGATTTAGCCTCATA – 3'

Oligonucleotide 2 (12-mer):

5' - TATGAGGCTAAA – 3'

Ligation was allowed to proceed overnight at 14^0C. After ligation, cDNA was size fractionated to get rid of small cDNA fragments (<500 bp) through a S400 sephacryl spin column according to the manufacturer's protocol (Promega).

3.3.3 PCR amplification of cDNA

PCR Amplification of cDNA was performed according to the method described by Carson and Botha (2000), with a few modifications. Ligated size fractionated cDNA (Section 3.3.2), was PCR amplified using oligonucleotide 1 as the primer. The final reaction conditions were as follows: 1X Taq DNA Polymerase buffer (50 mM KCl; 10 mM Tris-HCl, pH 9; 0.1% (v/v) Triton X 100); 600ng of oligonucleotide; 15ng of ds cDNA template; 100 μM of each dideoxynucleotide triphosphate (dNTPs); 3.6 mM $MgCl_2$; 1 unit of Taq DNA polymerase.

The PCR amplification was performed in a thermal cycler under the following conditions: 1 cycle of 73^0C, 2 min, to "melt" the short oligonucleotide off, followed by 35 cycles of 94^0C, 0.8 min; 68^0C, 1.1 min; and 73^0C, 3 min. An aliquot of each amplified cDNA sample was then analysed in a 1.5% (w/v) agarose gel to confirm success of the amplification reaction. The remaining PCR product was purified through a QIAquick-spin column.

3.3.4 Library construction

For library construction, 2μg of the amplified (Section 3.3.3) cDNA from the core and peripheral tissues of each internode were pooled into a total of 4μg and then ethanol precipitated. The cDNA was digested overnight with 30 units EcoRI at 37^0C. After heat inactivation of the enzyme at 70^0C for 15 min, the digested cDNA products were purified through a PCR cleanup column (Qiagen) and concentrated by ethanol precipitation. The cDNA pellet was then resuspended in 5μl of TE buffer (10 mM Tris-HCl; 1 mM EDTA, pH 8.0) and the concentration was determined by fluorometry (Dynaquant for DNA). The cDNA fragments were cloned into the EcoRI site of the λ Zap II phage (Stratagene) and the library was packaged using a GIGAPACK GOLD packaging extract (Gigapack II, Stratagene).

To determine the titre of cDNA libraries, aliquots of the constructed cDNA libraries were plated out separately onto solid NZY medium for blue-white selection, according to manufacturer's instructions (Stratagene). The core of primary differential plaques from plates were expelled into 200µl of SM buffer (100 mM NaCl; 8 mM $MgSO_4.7H_20$; 20 mM Tris-HCl, pH 7.5; 0.01% (w/v) gelatin) in 1.5ml tubes using the narrow ends of sterile pasteur pipettes. The phages were mixed in the buffer by vortexing thoroughly.

To examine the initial sizes of individual recombinant phages, 1µl of phage suspension was PCR amplified (20-30ng of DNA) using M_{13} reverse and T_7 primers. The final PCR mix conditions were as follows: 1 unit Taq DNA polymerase; 1X Taq DNA polymerase buffer; 3.6 mM $MgCl_2$; 200 µM of all 4 dNTPs; 6 µM of each primer and 1µl of phage suspension. PCR amplification was performed in a thermal cycler under the following conditions: 1 cycle at 94^0C, 0.5 min, followed by 10 cycles of 94^0C, 0.5 min; 50^0C, 0.75 min; 72^0C, 0.75 min. This was followed by 25 cycles of 94^0C, 0.75 min; 45^0C, 0.5 min; 72^0C, 0.5 min; 1 cycle of 72^0C, 2 min.

The PCR products were then analysed in a 1.5% (w/v) agarose gel with two DNA molecular weight markers III and V (Roche). Primary cDNA libraries were then amplified in XL1-blue MRF' *E.coli* cells and the titre determined according to instructions in the manual (Stratagene). The amplified cDNA libraries were then stored at -80^0C until used.

3.4 Preparation and Screening of Expression Arrays

Individual pBluescript SK(-) phagemids, containing the cloned cDNA inserts of internode 10 cDNA library (Section 3.3), were obtained by *in vivo* mass excision from the λ Zap II vector according to the manufacturer's instructions (Stratagene). These phagemid clones were plated out at a low density onto solid Luria Bertani (LB) medium containing 50 $\mu g.ml^{-1}$ ampicillin and incubated overnight at 37^0C. Using a colony picker (V & P Scientific), 3840 randomly chosen clones were transferred to 96–well microtitre plates containing 200µl

LB with 0.2 $mg.ml^{-1}$ ampicillin in each well. Bacterial clones were incubated overnight without agitation at 37^0C.

For cDNA expression array preparation representative of an internode 10 cDNA library, bacterial clones were arrayed on 13 x 10 cm positively charged nylon membranes (Roche) in a 4 x 4 duplication pattern using a library copier according to manufacturer's instructions (V & P Scientific, Inc). To lyse the bacterial cells, cDNA array filters were incubated for 5 min in 500 mM NaCl; 500 mM NaOH, neutralised for 5 min with 1.5 M NaCl; 1.5 M Tris and left to air dry for 1h. The DNA was then denatured using 400 mM NaOH, neutralised using 5X SSPE (100 mM sodium hydrogen phosphate; 1.8 M sodium chloride; 10 mM EDTA, pH 7.4) and left to air dry. This was followed by the UV crosslinking of the cDNA to the nylon membrane for 2.5 min at 120 $mJ.cm^{-1}$.

Membranes were hybridised with ^{33}P–labelled first strand cDNA fragments. The probes were prepared by reverse transcription of 1μg mRNA from 7C, 7P and 10P (5mm inner core (C) and 5mm outer peripheral (P) tissues of internode 7 and 5mm outer peripheral tissues of internode 10), according to the method of Sambrook *et al.* (1989), using Expand ™ Reverse Transcriptase (Roche) and with the addition of 10 μM ddCTP (Decraene *et al.*, 1999).

The cDNA array filters were prehybridised in a container at 65^0C for 6–18 h in 100ml hybridisation buffer (500 mM sodium phosphate; 7% (w/v) SDS; 1 mM EDTA; 1 $\mu g.ml^{-1}$ salmon sperm DNA). For hybridisation, fresh hybridisation buffer containing 10^7 cpm of the purified, denatured probe was added. Hybridisations were carried out at 65^0C overnight.

The cDNA filters were washed twice at room temperature for 20 min in 1x SSC (150 mM NaCl; 15 mM sodium citrate), 0.1% (w/v) SDS. A final wash at 65^0C for 20 min in 0.5x SSC, 0.1% (w/v) SDS was performed. After the washing steps, each cDNA-hybridised filter was wrapped in plastic wrap to avoid drying and exposed to a Super Resolution cyclone Phosphor Screen

(Packard) for 4–16 h. Data was captured and analysed using the phosphor imager and analysis system. Prior to the following hybridisation experiment radioactive probes were stripped from the cDNA nylon filters by washing twice in a solution containing 100 mM NaOH, 10 mM EDTA and 0.1% (w/v) SDS, with shaking for 10 min at room temperature. Thereafter, the membranes were rinsed in deionized water for 1 min, followed by a final wash in 5x SSPE for 10 min.

Differentially hybridised signals were selected in an empirical fashion by comparing the expression profile of hybridised cDNA targets on the arrays of the three independent hybridisation experiments.

3.5 Sequence Analysis of Differential Clones

Thirty selected cDNA clones were sequenced (ABI PRISM® dye terminator cycle sequencing) using the ready reaction kit with AmpliTaq® DNA polymerase (The Perkin Elmer Corporation, Norwalk, USA). cDNA fragments were sequenced from both ends using T_7 promoter primer and the M_{13} reverse primer. The cDNA sequences were edited to discard vector / linker and primer sequences.

The sequences were compared to the nucleotide (Non–Redundant and ESTs) and protein sequence databases by using the BLAST N and BLAST X internet server of the NCBI (blast@ncbi.nlm.nih.gov). cDNA sequences with matches to the same database accessions were compared to each other using DNASIS® for Windows® Sequence Analysis software (Hitachi Software Engineering Co., Ltd) to identify overlapping clones.

3.6 Northern Blot Analysis

Thirty microgram of total RNA from root, leafroll, young culm (internode 3), peripheral and core tissues of mature culm (internodes 7 and 10) were separated in a 1.2% (w/v) agarose gel. These samples were denatured by adding 50% (v/v) formaldehyde and 1x MOPS buffer (200 mM 3-[N –

morpholino] propanesulphonic acid; 50 mM NaOAC; 5 mM EDTA) and heating to 65^0C for 10 min before loading.

The 0.24–9.5 kb RNA ladder (Gibco BRL) was treated in a similar way as the samples. After separation in the gel, staining was carried out in ethidium bromide (0.1 $mg.ml^{-1}$). The 28S and 18S rRNA fragments were verified using the RNA ladder, which was then cut off from the gel. The gel was then destained in 30ml of deionised H_20. The samples of interest were then transferred overnight to a positively charged nylon membrane (Roche) by upward capillary blotting in 10x SSC. The RNA was linked to nylon membrane by exposing it to UV light for 2.5 min at 120 $mJ.cm^{-1}$.

Prehybridisation and hybridisation was performed in ULTRAhyb™ buffer (Ambion) according to the manufacturer's instructions. Prior to probe synthesis, the selected cDNA fragments were amplified as reported in Section 3.3.5. The cDNA fragments were then labelled using the Prime-It II random primer labeling kit (Stratagene), in the presence of [α-^{32}P] dCTP (Amersham; 3000 Ci $mmol^{-1}$). Membranes were exposed to Phosphor screens and data scored as described in section 3.4.

CHAPTER 4
Results

4.1 RNA Extractions and Poly $[A]^+$ Extraction

One of the main goals of this study was to develop a reliable protocol for the isolation of high quality RNA from the mature culm tissues of the sugarcane, to be used in cDNA synthesis and northern analysis.

High quality total cellular RNA from sugarcane culm tissues were successfully extracted with both the perchlorate and phenol based methods. These procedures were based on the ability of organic molecules to denature and precipitate proteins in solution while not affecting the solubility of the RNA (Dellaporta *et al.,* 1983). The integrity of total RNA was assessed by the presence of the two sharply defined ethidium bromide stained bands representing the ribosomal 28 and 18 RNA subunits (Figures 4.1 and 4.2).

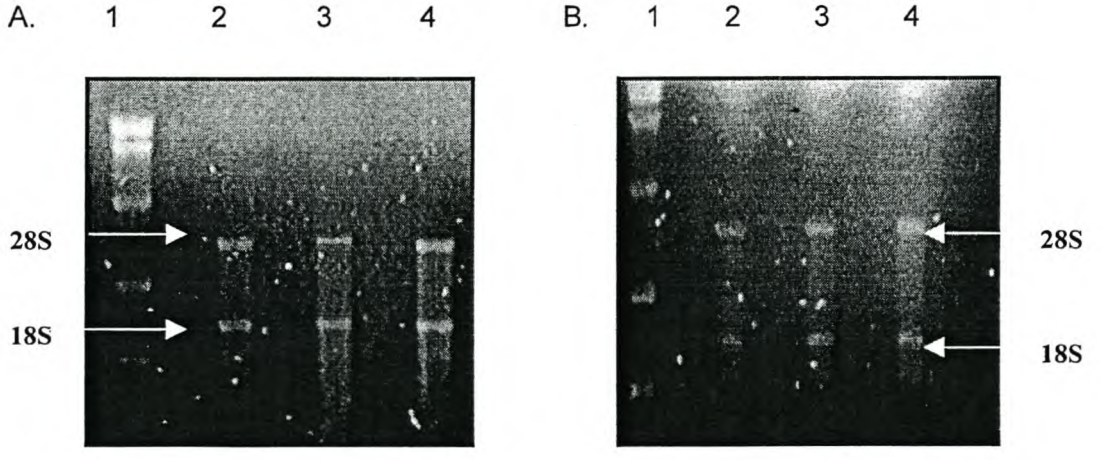


Figure 4.1 Gel electrophoretic analysis of representative samples of RNA extracted using the perchlorate method. RNA in samples (1μg/lane) was resolved in a formaldehyde containing 1.5% (w/v) agarose gel.
Lane 1: 0.24 – 9.5 kb RNA ladder (GibcoBRL); Lanes 2, 3 and 4: Total RNA from internodes 2, 7 and 10 of the sugarcane tissues respectively. A; total RNA from core tissues and B; total RNA from peripheral tissues. Distinct 28S and 18S ribosomal fragments are indicated with arrows.

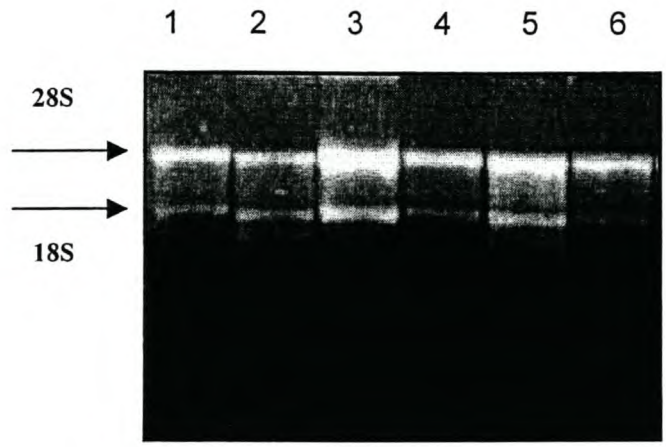


Figure 4.2 Gel electrophoretic analysis of representative samples of RNA extracted using the phenol based method. RNA in samples (1μg/lane) was resolved in a formaldehyde containing 1.5% (w/v) agarose gel.
Lanes 1 – 2: core and peripheral RNA samples of internode 2; Lanes 3 – 4: core and peripheral RNA samples of internode 7; Lanes 5 – 6: core and peripheral RNA samples of internode 10. Intact 28S and 18S ribosomal fragments are indicated with arrows.

The presence of a faint smear in the gel photographs was probably indicative of the presence of mRNA. Thus, agarose gel electrophoresis of the isolated RNA confirmed the undegraded nature of the samples.

To determine sample purity of all RNA preparations, the ratio of absorbance values determined at 260 nm and 280 nm ranged between 1.7 – 2.0 (Table 4.1). This indicated that the RNA was without protein contamination, since the A_{260}/A_{280} of uncontaminated RNA is approximately 2.

However, to select for the most efficient RNA extraction method applicable to the sugarcane culm tissues at various developmental stages, the phenol and perchlorate methods were compared in terms of their RNA yields per fresh gram tissue. As shown in Table 4.1, the phenol based method produced threefold or more quantites of RNA as compared to the yields obtained with the perchlorate method. Another interesting observation made in the RNA yield comparative analysis was the decreasing quantity of RNA obtained per

fresh gram of tissue with both methods from both the core and peripheral tissues of internode 2 to internode 10 (Table 4.1).

Table 4.1 Absorbance characteristics and yields of RNA extracted from sugarcane tissues by two extraction methods. All values are the mean of three independent isolation experiments.

	Perchlorate Method			Phenol Based Method		
Sample I_XY	A_{260}	A_{260}/A_{280}	RNA yield $\mu g.g^{-1}$ tissue	A_{260}	A_{260}/A_{280}	RNA yield $\mu g.g^{-1}$ tissue
I_2C	0.198	1.80	24.8	0.560	1.98	100
I_2P	0.291	1.96	36.38	0.515	1.99	92
I_7C	0.175	1.80	21.88	0.487	2.00	87
I_7P	0.169	1.70	21.13	0.398	1.97	71
$I_{10}C$	0.155	1.70	19.38	0.409	2.00	73
$I_{10}P$	0.160	1.80	20.00	0.336	1.98	60

I_XY: I_X – Internode Developmental Stages 2, 7 and 10
Y – Tissue Type, C for Core and P for Peripheral

Following poly $[A]^+$ RNA selection, the percentage yield of mRNA obtained from the total RNA used per each tissue sample was determined and ranged between 0.71-0.89 % as summarised in Table 4.2. This indicated that mRNA was successfully extracted since a reasonable estimate of poly $[A]^+$ RNA concentration was reported to be 1% of the total RNA used in its preparation (Croy *et al.*, 1993).

Table 4.2 Percentage yield of mRNA isolated from individual sugarcane internodal RNA samples.

	mRNA Samples					
I_XY	I_2C	I_2P	I_7C	I_7P	$I_{10}C$	$I_{10}P$
% Yield	0.89	0.79	0.71	0.76	0.80	0.75

I_XY: I_X – Internode Developmental Stages 2, 7 and 10
Y – Tissue Type, C for Core and P for Peripheral

4.2 Characteristics of the Internode 7 and Internode 10 cDNA Libraries

Oligo (dT) primed cDNA was successfully amplified (Figure 4.3). Thus, the success of the polymerase chain reaction (PCR) based cDNA amplification confirmed both efficient cDNA synthesis and successful adaptor ligations. It also generated a smear of cDNA fragments, which was indicative of the representative message profile in the core and peripheral tissues of the sugarcane culm at three developmental stages. The size of cDNA fragments larger than 0.5 kb, suitable for cloning, was confirmed (Figure 4.3) and were pooled to construct the representative cDNA libraries of internodes 7 and 10.

Blue-white plaque selection following incubation of an aliquot of each library in the presence of X–gal (5–bromo–4–chloro–3–indolyl-β-D–galactopyranoside) and IPTG (isopropyl-β-D-thiogalactoside), revealed that 95% of the internode 7 cDNA library clones and 97% of the internode 10 cDNA library clones were recombinant. The library titers were 2.46 x 10^6 pfu.μg^{-1} and 4.2 x 10^6 pfu.μg^{-1} cDNA, respectively.

The insert sizes of the randomly selected recombinants ranged between 0.5 – 2.0 kb and 0.5 – 3.0 kb for internodes 7 and 10 cDNA libraries, respectively (Figure 4.4 and Table 4.3).

Table 4.3 Analysis of cDNA libraries representative of the sugarcane culm at two developmental stages.

cDNA Libraries	Internodes 7	10
Titre Determination pfu.μg^{-1} cDNA	2.4 x 10^6	4.2 x 10^6
Average Insert Sizes (kb)	0.5 – 2.0	0.5 – 3.0

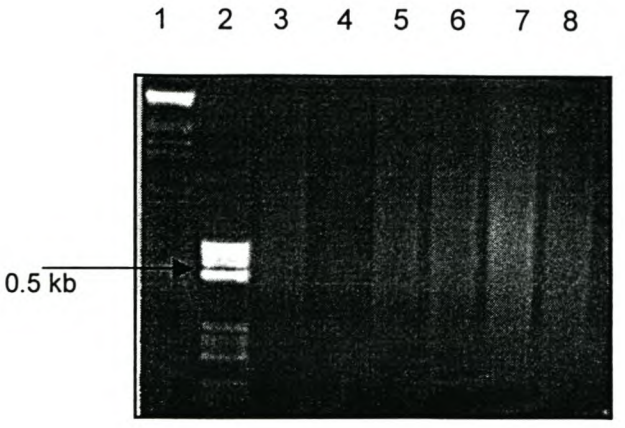


Figure 4.3 Gel electrophoretic analysis of amplified cDNA fragments prepared for cDNA library construction. cDNA in samples (0.25μg/lane) were resolved in a 1.5% (w/v) agarose gel. Lanes 1 and 2: Molecular weight markers III and V; Lanes 3 and 4: core and peripheral samples of internode 2; Lanes 5 and 6: core and peripheral samples of internode 7; Lanes 7 and 8: core and peripheral samples of internode 10. The arrows indicate amplified cDNA fragments larger than 0.5 kb.

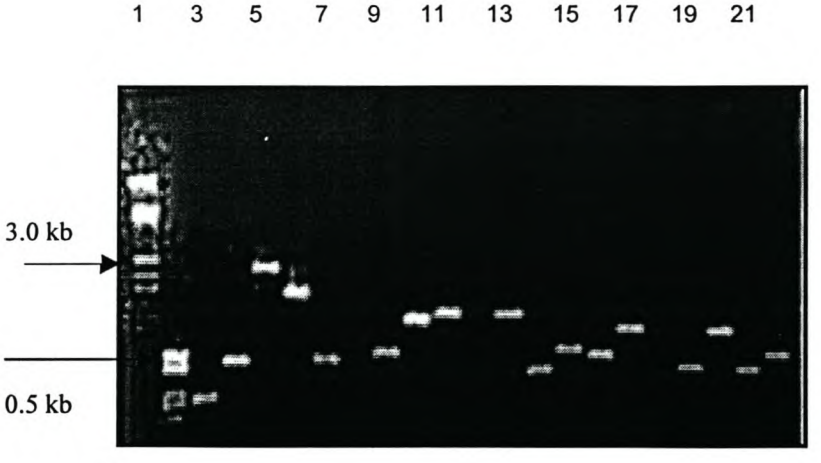


Figure 4.4 Characterisation of the sugarcane internode 10 cDNA library inserts. PCR amplification of randomly selected cloned cDNA fragments was primed using T7 and pUC18/M13 reverse primers. Lanes 1 and 2: Molecular weight markers III and V; Lane 3: amplification product of the phagemid with no insert; Lanes 4 – 22: amplification products of the phagemids with inserts of various sizes. The arrows indicate the insert sizes of the randomly selected recombinants in the range of 0.5 – 3.0 kb.

4.3 Screening for Potential Differentially Expressed Clones using cDNA Expression Arrays

The differential screening of the internode 10 cDNA library was the first step towards the identification of genes which were differentially expressed in the core and peripheral tissues of the mature sugarcane culm. The array filters were screened in parallel with radioactively labelled cDNA representatives of the mRNA populations derived from three specific tissues of the sugarcane culm (see section 3.4). The specific activity of the probe prepared from core tissues of internode 7, peripheral tissues of internode 7 and peripheral tissues of internode 10 were 4.07 x 10^7, 7.45 x 10^7 and 6.25 x 10^8 cpm. μg^{-1} DNA, respectively. Image analysis for each screened array was then performed through the extraction of data from phosphor image representations of hybridisation signals. One of the ten membranes of the array screening is presented in Figure 4.5 as a representative image example.

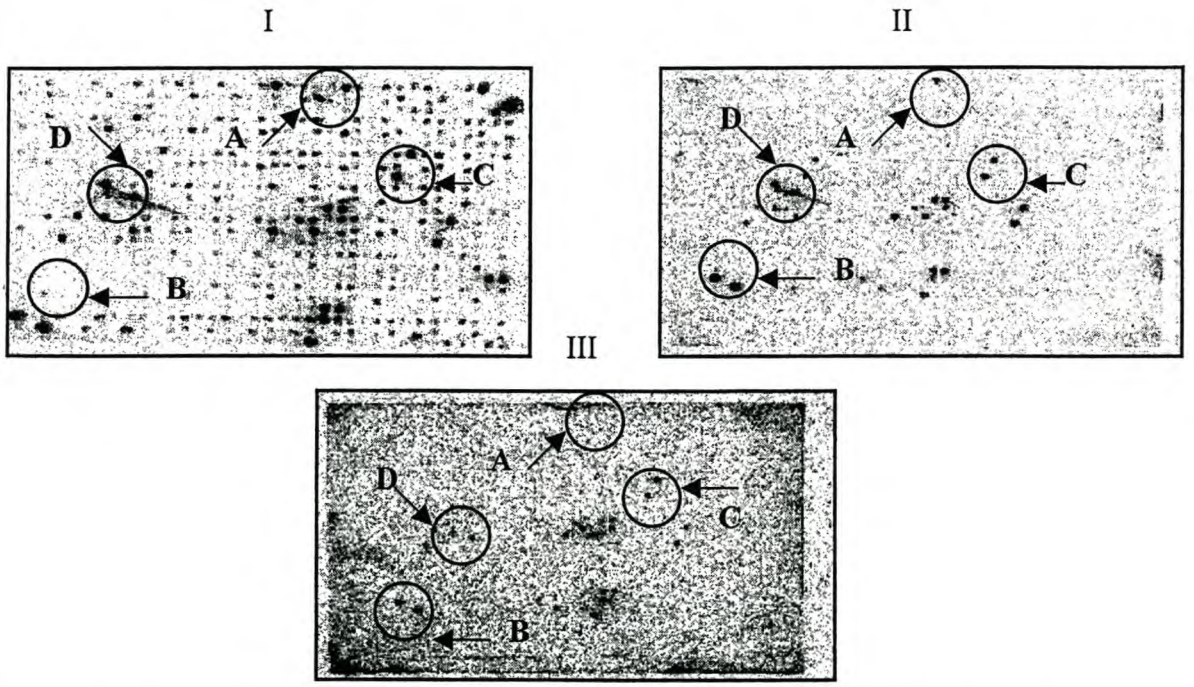


Figure 4.5 Expression profile of 384 genes from sugarcane internode 10. The arrays were screened in parallel with cDNA populations derived from internode 7 core tissues (I), internode 7 peripheral tissues (II) and internode 10 peripheral tissues (III) of the mature sugarcane culm. A, B C and D: core specific, peripheral specific, selectively expressed and stage specific differentially hybridised signals, respectively.

Visual inspection of the images firstly revealed the success of the three independent hybridisation experiments with the detection of sharply defined radioactive signals at several locations on the arrays. The radioactive signals represented hybridisation of the arrayed DNA to its complementary molecule derived from the probe used to screen the array. Hence, hybridisation signals and their respective intensities were indicative of the relative transcript abundance of the arrayed cDNA in the core and peripheral tissues of internode 7 as well as in the peripheral tissues of internode 10.

Variation in background radioactivity was also observed on the arrays screened with probes generated from the three tissues. However, it had no effect on the identification of sharply defined hybridisation signals such as A, B, C and D in Figure 4.5, of which the intensities exceeded that of the local background radioactivity. Background radioactivity may be contributed due to the hybridisation kinetics of the different probes (Duggan *et al.*,1999). Overall, clear detection of strong hybridised signals with relatively lower background radioactivity confirmed the good quality of the probes and the efficiency of the stringent hybridisation and wash conditions used for screening abundantly expressed genes.

To assess the reproducibility of the arraying and hybridisation process, the intensity of the hybridisation signals for the duplicates of each clone on the same filter were compared. The signals were reproducible with the detection of defined duplicated spots with no difference in signal intensity. This observation eliminated variable and therefore less reliable gene expression data for further analysis. It also indicated that there was no significant variation in the quantity of DNA spotted during array fabrication, thus making the technique sensitive for monitoring gene expression levels.

Given the ability and sensitivity of the array technology for monitoring gene expression, the relative transcript abundance of each arrayed cDNA in the three labelled representations of the cellular cDNA pools, was determined. This was achieved by visually examining the hybridisation pattern at individual element level. Analysis of hybridisation data revealed that the three labelled

representations of the cellular cDNA pools did not hybridise to each of the 3840 immobilised cDNA clones on the arrays. The percentages of the total number of clones producing strong signals representing expression of highly abundant genes were low. These were 3.5, 2.6 and 2.5% for arrays probed with labelled cDNA representations of the core tissues of internode 7, peripheral tissues of internode 7 and peripheral tissues of internode 10 of the mature sugarcane culm, respectively. However, it became evident that the level of gene expression was higher in the core tissues rather than the peripheral tissues of the mature sugarcane culm.

The expression profiles of the hybridised cDNA targets on the arrays were also compared. A total number of 50 clones from the internode 10 cDNA library were expressed at equal levels on all the probed arrays which indicated that they were constitutively expressed in the mature tissues of the sugarcane culm. However, the criteria used to select for differentially expressed genes specific to a tissue type of the sugarcane culm was based on identifying arrayed cDNAs that hybridised exclusively to one of the three probes. Hence, cDNAs displaying a differential pattern of hybridisation signal were identified as indicated by the letters A, B C and D in Figure 4.5.

As an example, arrayed cDNA clone denoted A, was preferentially expressed in the core tissues of internode 7 (I) and showed no expression in the peripheral tissues of the sugarcane culm (II and III). In contrast, cDNA clone denoted B was selected as up-regulated in the peripheral tissues of the mature sugarcane culm (II and III), while no expression was detected in the core tissues (I). Clone C is an example of a cDNA element that was expressed most abundantly in the core tissues of the sugarcane internode 7 (I), with a decreased pattern of expression in the peripheral tissues of the same (II) and older developmental stage of the sugarcane culm (III). Furthermore, arrayed cDNA clone denoted D was preferentially expressed in both the core and peripheral tissues of internode 7 (I and II), while low expression was observed in the peripheral tissues of internode 10 (III). Thus, comparisons of the gene expression in the three independent hybridisation experiments led to the identification of differentially expressed genes not only

specific to the core and peripheral tissues of the culm (gene expression pattern represented by A and B respectively), but genes that were up-regulated and down-regulated in the mature sugarcane internodes (of which a decrease in expression level towards the more mature internode 10 is represented by C). In addition, internode stage specific clones were also identified (of which specific expression to internode 7 is represented by D). The analysis of gene expression in the mature sugarcane culm using cDNA expression arrays has led to the identification of 125 differentially expressed genes as summarised in Table 4.4. This number of genes was more than sufficient for the purpose of this study.

Table 4.4 The number of differentially expressed sugarcane genes identified by cDNA expression arrays representative of the internode 10 cDNA library.

Type of Expressed Clone	Number of Clones Collected
Peripheral Specific	15
Core Specific	45
Selectively Expressed	50
Internode 7 stage specific	10
Internode 10 stage specific	5

4.4 Sequence Analysis of Differentially Expressed Clones

The sequences of 30 differentially expressed clones identified in section 4.4 were successfully analysed. This group of randomly chosen clones was represented by 9 core specific, 5 peripheral specific and 16 selectively expressed gene sequences derived from the mature culm. Manual editing of vector sequences of these 30 differentially expressed clones revealed the presence of multiple adaptor sequences within the insert fragment of each cloned cDNA. This result indicated that the insert sizes of the cDNA libraries were previously overestimated (Figure 4.4). After removing the vector and adaptor sequences, a total number of 91 unique sequences from the mature culm tissues were obtained. These were 27, 14 and 50 core specific, peripheral specific and selectively expressed gene sequences of the culm, respectively. Expressed Sequence Tags (ESTs) were then established for the

differentially expressed cDNAs based on sequence similarity to known proteins.

Deduced amino acid sequence homology between a sugarcane EST and a known sequence homology was deemed significant if the BLASTX Point Acceptable Mutation (PAM) 120 similarity score was greater than 80 (Altschul *et al.*, 1990). Database search results revealed that sequence homology was not obtained for any of the cDNAs displaying core and peripheral tissue specific expression in the mature culm. This suggested that the transcripts may encode proteins of unknown function and may represent new genes not only in sugarcane but also in other organisms. Of the 50 cDNAs which exhibited selective expression in the mature culm (cDNAs showing stronger expression in the core than in the peripheral tissues), only 33 cDNA sequences exhibited homologies to gene sequences from other organisms, while the remaining 17 gene sequences remained unidentified. However, only 10 out of the 33 identified sequences had a PAM 120 similarity score of over 80 and were considered homologous proteins for the clones (Table 4.5). The other 23 sequences did not show significant homology to previously identified genes in the databases (i.e. similarity scores below 80) and thus were putatively identified on the basis of sequence similarity only (Table 4.5).

All sugarcane cDNAs showed primary homology to gene sequences from organisms other than sugarcane, except for the gene homologous to the polyubiquitin protein and drought inducible 22kD protein (represented by clone SCLO 2_b and SCLO 4_b in Table 4.5, respectively). The homologous genes were from plants such as barley, *Oryza sativa* and *Zea mays*. This suggests that these cDNAs may not only be plant-specific but may also be specific to monocotyledonous plant species. Furthermore, several cDNAs encoding highly abundant proteins were isolated more than once and this also includes multiple copies of the same cDNA sequence encoding a specific gene. This was observed for cDNA encoding the jacalin homolog (15x), HSP70 (12x), smHSPs (10x), drought inducible 22kD protein (16x), water stressed induced protein (9x), probable Beta-Amylase (7x), SNF1 like protein kinase (8x), hydroxyproline rich glycoprotein (12x), polyubiquitin (7x), sucrose

synthase (8x) and translational initiation factor (7x). Nevertheless, it has to be emphasised that the presented functional identities of the genes in Table 4.5 are only indications of the possible function of the corresponding protein. A true biological function has to be demonstrated by biochemical and genetic studies. However, the frequency of cDNA clones corresponding to the same gene can provide an estimate of steady-state transcript levels in a particular tissue (Cooke *et al.,* 1996). Hence, genes present in multiple copies in this study could represent an increase in the expression of the corresponding genes in the maturing sugarcane culm. In contrast, the remaining cDNAs were represented by 2 to 3 copies, probably indicative of the expression of relatively less abundant classes of genes.

Furthermore, a consensus sequence of an 1080 bp fragment was generated for cDNAs showing homology to the jacalin homolog through an identical alignment with approximately 200 bp overlap. However, no other consensus sequences for the remaining putatively identified genes, were derived; firstly due to the fragments being too small (in the range of 100–150 bp) with no sequence overlaps; secondly, frequently detected genes had identical sequences. Since 64% of the total number of cDNAs remained unidentified, clones with no database match were used to search the EST database to identify possible matches to previously identified ESTs. Results indicated that although all clones did show some sequence similarity to a known EST, in most cases the sequence homology was not significant.

Differentially expressed genes were assigned putative identities and their respective functions in the core tissues of the mature culm were putatively characterised according to the proposed function of their homologues in other organisms. Many genes were found to show sequence similarity to a wide variety of genes associated with the maintenance and control of cellular metabolism.

These included cDNAs that were homologous to genes associated with the replication, transcriptional and translational apparatus (eg. DNA replication licensing factor, probable reverse transcriptase, translational initiation and

elongation factors) in the assembly of ribosomal structure (eg. ribosomal proteins L16 and S3) and in signal transduction processes (eg. protein kinase C inhibitor, SNF1 like protein kinase). Also identified were genes associated with degradation processes (eg. polyubiquitin), with osmotic adjustment (eg. water–stressed protein), and with cytoskeletal components (eg. structural polyprotein C, hydroxyproline rich glycoprotein, actin depolymerising factor 3). Only one cDNA exhibited sequence homology to gene associated with sucrose metabolism (sucrose synthase). Most of the remaining genes could be categorised according to putative roles in stress responses (eg. drought inducible 22 kD protein, jacalin homolog and heat shock protein).

Table 4.5 Differentially expressed genes in the mature sugarcane culm with sequence homology to known plant genes.

Clo No_X	Origin	Confirmed Northerns	Fragment Size (bp)	Maximum Score	Putative Identity	Organism
SCLO 16_a	7C	Yes	790	64.7		Barley
SCLO 22_a	7C	Yes	414		Jacalin homolg - barley	
SCLO 25_a	7C	Yes	265			
ConSQ_1			1080		Jacalin homolg - barley	Barley
SCLO 23_a	7C	Yes	283	34.3	Putative translational initiation factor EIF – 2B alpha subunit	*Arabidopsis thaliana*
SCLO 22_c	7C	Yes	614	70.9	Putative gag – pol polyprotein, 3' partial	*Oryza sativa*
SCLO 23_c	7C	Yes	186	29.3	Modular polyketide synthase	*Gossypium tirsutum*
SCLO 22_d	7C	Yes	135	63.5	Putative DNA replication licensing factor Mcm5	*Arabidopsis thaliana*
SCLO 23_d	7C	Yes	129	69.3	Probable reverse transcriptase	*Oryza sativa*
SCLO 23_e	7C	Yes	138	34.7	NADH dehydrogenase I chain G	*Gossypium tirsutum*
SCLO 23_f	7C	Yes	301	34.7	Serine/threonine protein kinase	*Triticum aestivum*

SCLO 1_a	7C	No	452	85.1	Putative protein	*Arabidopsis thaliana*
SCLO 1_b	7C	No	204	121	ZB4 - Maize – 14KD Zinc Binding Protein (Protein Kinase C Inhibitor	*Zea mays*
SCLO 2_a	7C	No	268	54	Hydroxyproline rich glycoprotein	*C. roseus*
SCLO 2_b	7C	No	526	162	Polyubiquitin protein	*Saccharum hybrid* cultivar
SCLO 3_a	7C	No	403	30	Probable Beta Amylase	*Arabidopsis thaliana*
SCLO 4_a	7C	No	350	211	Putative sucrose synthase	*Arabidopsis thaliana*
SCLO 4_b	7C	No	472	63.2	Drought inducible 22KD protein	*Saccharum officinarum*
SCLO 4_C	7C	No	391	394	*Zea mays* ABA and ripening inducible protein	*Zea mays*
SCLO 5_C	7C	No	452	89	Water stressed induced protein	*Arabidopsis thaliana*
SCLO 6_a	7C	No	441	82	Similar to HSP70 protein	*Zea Mays*
SCLO 7_b	7C	No	388	249	ADP – ribosylation factor 1	*Oryza sativa*
SCLO 8_a	7C	No	155	33.1	22 kDa Induced kafirin cluster protein	Sorghum
SCLO 8_b	7C	No	359	235	Actin depolymerising factor 3 Structured polyprotein C,	*Zea mays*
SCLO 8_c	7C	No	147	29.3	contains coat polyprotein C; spike glycoproteins E_3, E_2 and E_1 6KD polypeptide	*Zea mays*
SCLO 9_a	7C	No	453	37.4	SNF1 like protein kinase	*Arabidopsis thaliana*
SCLO 9_b	10P	No	259	40	Major Latex protein type 2	*Arabidopsis thaliana*

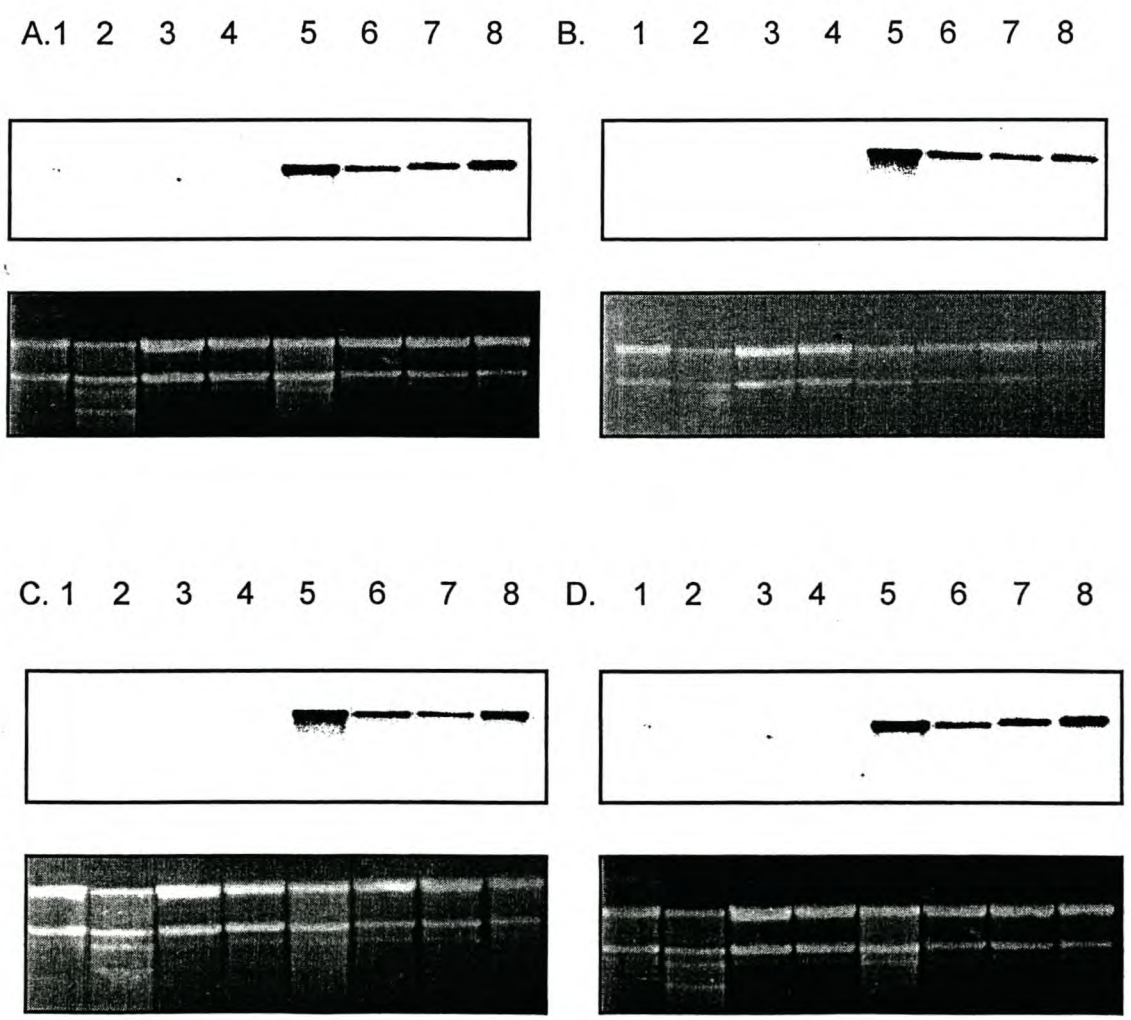
SCLO 9_c	7C	No	185	34.2	Hypothetical RNA binding protein	*Arabidopsis thaliana*
SCLO 10_a	7C	No	185	40	Translation elongation factor 1 alpha	Wheat
SCLO 11_a	7C	No	827	196	*Zea mays* NCS 3 mutant DNA for mitochondrial L16 – like and S3 – like ribosomal proteins	*Zea mays*
SCLO 12_a	7C	No	678	198	*Zea mays* high mobility group I/Y – 2 mRNA	*Zea mays*
SCLO 13_a	7C	No	272	30.4	Plasma membrane intrinsic protein	*Arabidopsis thaliana*
SCLO 14_a	7C	No	150	40	NADP – dependent oxireductase like protein	*Zea mays*
SCLO 15_a	7C	No	265	50	SmHSP	*Arabidopsis thaliana*
SCLO 16_a	10P	No	289	299	Putative Mitochondrial Retroelement protein	*Oryza sativa*

The heading CLOX_a designates X for clone identity and a, representative of the first fragment within the multiple cloned insert. Origins of clones showing maximum expression on the expression arrays are indicated where 7C and 10P represent core tissues of internode 7 and peripheral tissues of internode 10 respectively. The validity of the results obtained with the arrays (i.e. identification of potential differentially expressed genes in the mature culm) were further evaluated on northern blots for clones as indicated. The headings fragment size indicate the length of analysed sequences of the sugarcane genes in base pairs; Maximum Score represents the BLASTX PAM120 highest homology score to known plant sequences in the international databases. Organism refers to the source of the protein with the highest sequence similarity to the sugarcane genes. Analysed sequences and consensus sequence of the jacalin homolog barley (ConSQ_1) are provided in Appendix A.

4.5 Northern Analysis of Potential Differentially Expressed Genes

In Figure 4.6, the expression patterns of five potential differentially expressed sequences, which were identified from array screening (Section 4.3) are displayed. Clones SCLO 14, SCLO 16, SCLO 22, SCLO 23 and SCLO 25 showed differential gene expression in the core and peripheral tissues of internodes 7 and 10 only. No gene expression was observed in root, leaf, leafroll and internode 3. The specific activity of their respective probes were (7.9, 1.9, 5.2, 5.6 and 4.5) x 10^9 cpm $\mu g.DNA^{-1}$, respectively.

Northern analysis was also performed for 10 differentially expressed clones that showed specific core and peripheral expression in the tissues of the mature sugarcane culm according to results obtained from the screened cDNA expression arrays of internode 10 cDNA library. However, no expression of those genes could be detected under the hybridisation conditions used for the northern analysis.



E. 1 2 3 4 5 6 7 8

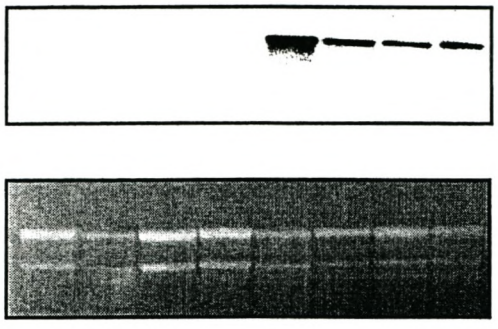


Figure 4.6 Northern hybridisation analysis for clones A) S CLO 14, B) S CLO 16, C) S CLO 22, D) S CLO 23 and E) S CLO 25. RNA gel photographs are included to confirm equal loading of total RNA. Lanes: 30 µg total RNA isolated from 1) root, 2) leaf, 3) leafroll, 4) internode 3, 5) internode 7 core, 6) internode 7 peripheral, 7) internode 10 core and 8) internode 10 peripheral.

CHAPTER 5
Discussion

5.1 RNA Analysis in the Core and Peripheral Tissues of the Sugarcane Culm

Extraction of good quality total cellular RNA from core and peripheral tissues of the mature culm, using the phenol based method, led to the successful development of good quality cDNA libraries representative of the mature culm. Despite the tissues of the sugarcane being composed of storage parenchyma and vascular bundles surrounded by an increased amount of cell wall lignification down the culm (Jacobsen *et al.,* 1992), the phenol method effectively disrupted the tissues, producing significant RNA yields for this study. In contrast, this method produced poor RNA yields from grape berries, which contained high levels of sugars and phenolic compounds (Tesniere and Vayda, 1991).

However, the decrease in RNA yield observed per gram of fresh tissue from the various core and peripheral tissues down the sugarcane culm, was indicative of tissues being less metabolically active in older developmental stages as compared to the younger developmental stages. For example, metabolic activities such as structural polysaccharide synthesis, storage polymer synthesis and respiration have been reported to be higher in younger sugarcane stem internodal stages with an altered pattern observed in the older internodal stage (Whittaker and Botha, 1997). Furthermore, the RNA quality and integrity obtained were very crucial for this study since enriching for mRNA from each specific tissue of the sugarcane culm should provide a message profile of the genes expressed. Thus, the presence of all the transcripts expressed in the core and peripheral tissues of the mature culm ensures the production of a representative cDNA library. However, the relatively low and constant level of mRNA yield obtained from the various sugarcane tissues did not correlate with the decreased pattern of RNA yields down the mature culm. This was probably indicative of only a subset of sugarcane genes being expressed in the regulation of culm metabolism.

The expression of many genes associated with lipid metabolism (eg. acyl-Co-binding protein), cell division and growth (eg. vacuolar H^+-ATPase), and RNA synthesis (eg. RNA polymerase I, II, III) was reported to be down-regulated in the culm tissues (Carson, 2001). In contrast, their level of gene expression increased in the sugarcane immature and mature leaf (Carson, 2001). Hence, in-depth molecular analyses are required to develop an understanding of the significance of the decreased gene activity of these genes towards the regulation of culm metabolism.

5.2 Significance of the cDNA Library Produced and Future Use

The internodes seven and ten cDNA libraries produced, with 95% and 97% of recombinancy respectively and a titer of 10^6 pfu.μg^{-1} cDNA, served as a good source of well defined sequences from the core and peripheral tissues of the mature sugarcane. The average insert sizes from 1.0 – 1.8 kb of the two cDNA libraries was an indication that full length cDNA molecules could be present. Consequently, a library containing full-length cDNA clones was considered to be an advantage for screening the complete range of cDNA sequences of the messages present in the two specific tissue types of sugarcane at late developmental stages. However, even though it could be argued that the titer in the range of 10^6 pfu.μg^{-1} cDNA was comparatively low relative to the complexity of the polyploid sugarcane genome, they were considered to be sufficiently representative for preliminary analysis of the abundantly expressed genes present. This was because the distribution of mRNA present in high abundance is believed to represent 22% of the total mRNA population, with a relative abundance of approximately 3,500 copies of each individual abundant mRNA transcript per typical eukaryotic cell (Zubay, 1993). Therefore, the relative frequency of cDNAs in the mature culm tissue types, reflecting genes belonging to the superabundant and abundant classes, were likely to be cloned. Such type of investigation has been performed and reported by Carson and Botha (2000), with leaf roll and maturing culm cDNA libraries of similar titre.

Through the application of Expressed Sequence Tag (EST) analysis in large scale cDNA sequencing projects for a variety of plants such as *Arabidopsis thaliana* (Newman *et al.*,1994) and rice (Sasaki *et al.*, 1994), putative clones for a wide range of gene products were obtained. Since a limited number of genes in the sugarcane culm have been identified, the cDNA libraries produced in this study could be very useful for similar projects in sugarcane. This would lead to the establishment of an EST database representative of the core and the peripheral tissues of mature sugarcane culm. Hence, availability of these specific DNA sequences would be beneficial due to the shortage of plant ESTs in the international database compared to ESTs derived from human and other animal species (Carson and Botha, 2000).

5.3 Differential Gene Expression in Sugarcane Mature Culm Monitored using cDNA Arrays

In this study, a procedure for the parallel monitoring of 3680 gene sequences of the mature sugarcane culm using the cDNA expression arrays was described. By randomly selecting clones from a late developmental stage cDNA library for expression analysis, it was anticipated that this would provide some indication of the morphological and physiological differences between the two tissue types of the maturing culm. As shown in this study, the differential screening was best suited at rapidly decreasing the large numbers of candidate genes in a highly specific manner, limiting the results of the screen to genes of which the expression levels differed in the tissue types of the mature culm. However, only a small percentage (3.3%) of the cDNAs were found to reflect differential levels of gene expression between the core and peripheral tissues of the two late internodal stages. This was in contrast to research conducted where comparisons between the expression patterns of genes in the leaf and mature culm tissues of sugarcane revealed a relatively higher number (up to 8.6 %) of differentially expressed genes (Carson, 2001). This may be due to the fact that the tissues used in this investigation were from the same organ, namely the mature culm and that Carson (2001) compared expression levels in leaf and culm tissues which are different organs. Furthermore, comparisons between the expression patterns

of genes in leaves, roots, flower buds and open flower in *Arabidopsis* revealed large numbers (up to 34%) of differentially expressed genes (Ruan *et al.*, 1998). Thus, it becomes evident that changes in gene expression increases as organs diversify, probably reflecting their morphological and physiological differences in the plant.

Moreover, this investigation was marked by the identification of 1.6% and 0.4% of the total number of cDNAs which displayed a tissue specific and stage specific pattern of expression in the mature culm, respectively. These were previously not detected in gene expression analysis of the leaf and maturing stages of the sugarcane culm tissues (Carson, 2001). However, the combination of enrichment of gene sequences through subtraction with the efficiency of array screening, which was normally used to detect very low abundance messages (0.02-0.05% of the total mRNA pool) (Aguan *et al.*, 1991), was required to detect tissue specific expression in the polyploid monocot. Thus, these observations firstly highlighted the high capacity of the cDNA array method for the sugarcane gene expression monitoring. Secondly, the cDNAs showing a differential pattern of expression in this study, were representative of genes that belonged to the abundant and superabundant classes. Similarly, highly abundant expressed genes (>0.5% of the total mRNA pool) were also identified in differential gene expression analysis in *Arabidopsis* and animal systems (DeRisi *et al.*, 1997; Ruan *et al.*, 1998).

Identification of groups of abundantly expressed tissue specific, stage specific and selectively expressed genes and their possible functions in the sugarcane mature culm did not provide an insight into their specific roles in the plant. Their respective differential gene expression patterns detected by the arrays were informative in the sense that they provided static information (that is in which tissues the genes are expressed) and dynamic information (that is how the expression pattern of one gene relates to another) (Duggan *et al.*, 1999), thus indicating the changes in the developmental program of the culm tissue types. Furthermore, genes with similar expression behaviour (for example, increasing or decreasing together under similar circumstances as observed in this study) are likely to be related functionally (Lockhart and Winzeler, 2000).

For example, differential cDNA expression data analysed from yeast cells showed that co-expressed genes dominated by more than 100 cytoplasmic ribosomal proteins and genes (such as elongation factors) involved in translation were all strongly down regulated in response to starvation and shock (Eisen *et al.*, 1998). In contrast, a distinctly different pattern was observed by a large cluster of co-ordinately expressed genes involved in mitochondrial ATP synthesis and electron transport. Similarly, groups of differentially expressed genes identified in this study can be used to establish a functional map (Brown and Botstein, 1999), representative of the patterns of gene expression acquired on a broad scale (overview of the choreography) and fine scale (gene by gene) in the mature culm tissues. The availability of such a map and information on the identity of their respective genes can be used to understand how the components work together to comprise functioning cells in the culm tissues of the polyploid monocot.

Furthermore, the great value and significance of transcript surveys in this study was accounted as follows; the increasing and decreasing pattern of gene expression down the culm as well as stage and tissue specific expression profiles observed, were indicative of the presence of specific sequence motifs in the promoter regions of the adjacent genes. This governs where and when the product was to be made and in what quantity (Lockhart and Winzeler, 2000). As an example, yeast studies showed that more than 50% of the genes that were transcribed in a cell cycle-specific manner and whose transcript abundance peaked in the G1 phase of the cell cycle, had an MCB (Mlu cell-cycle box) within 500 base pairs (bp) of their translational site (Tavazoie *et al.*, 1999).

The correlation between the presence of specific sequence motifs in promoter regions and gene expression patterns may be stronger than the correlation between functional categories and gene expression patterns (Lockhart and Winzeler, 2000). Hence, new *cis* regulatory elements (genomic sequence motifs that are over-represented in the genomic DNA in the vicinity of similarly behaving genes) may be revealed by examining classes of co-regulated genes, eliminating the need for a more conventional approach, using site-

directed mutagenesis ('promoter bashing') (Lockhart and Winzeler, 2000). For example, comparison of the *Vitis vinifera* hexose transporter 1 (*Vvht1)* promoter with the promoter of grape alcohol dehydrogenase, allowed the identification of a 15-bp consensus sequence, which suggested a possible co-regulation of the expression of these genes during the ripening of grape berry (Fillion *et al.*, 1999). Similarly, the differentially expressed gene fragments identified in this study can be used as potential research tools for tissue and stage specific promoter isolation from the mature culm, which is the long term goal of this study.

During sugarcane development, the rate of sucrose accumulation changes with internode expansion and maturation (Whittaker and Botha, 1997). In addition, physiological studies of key enzymes associated with sucrose metabolism in sugarcane such as sucrose synthase, sucrose phosphate synthase (SPS) and the various invertases (neutral, soluble and cell wall bound) have established that enzyme activities vary in internodes of differing maturity (Moore, 1995). Due to high levels of activity for these enzymes in the culm, it may be anticipated that transcripts for these genes, as well as others associated with sucrose metabolism, would be present at high levels. Thus, the identification of one hundred and twenty-five differentially expressed genes from the culm displaying a higher level of expression in the core rather than the peripheral tissues, was expected to represent candidate genes controlling sucrose metabolism.

However, EST analysis of a mature culm cDNA library showed that none of the high abundantly expressed transcripts were homologous to genes associated with sucrose metabolism (Carson and Botha, 2000). Similar observations were made for the highly abundant differentially expressed genes previously identified in the culm despite the high metabolic activity of enzymes involved in sucrose metabolism (Carson, 2001). Thus, the results presented here, points towards the possibility that the large numbers of genes associated with sucrose metabolism may belong to the medium and low abundance classes, which would not be detected with this procedure. Consequently, this was reflected by the large number of cDNAs on the filter

arrays that failed to reflect expression in any of the tissue types, due to them not being highly abundant transcripts. Thus, differentially expressed genes identified in this study could be representative of candidate genes regulating internodal expansion and culm maturation.

5.4 ESTs Analysis of the Potential Differentially Expressed Genes in the Mature Culm

The use of an EST approach was found to be a very efficient and successful method of identifying genes that are differentially expressed in the mature sugarcane culm (Ruan *et al.*, 1998). Since the cDNA library used was not a directional library, the orientations of the cDNA inserts were at random. This meant that it was not known from which end (5' or 3') the clones had been sequenced. However, to enhance the probability of obtaining sequence information of the coding region, selected differentially expressed clones were sequenced from both ends. However, the intricacy of the cDNA library methodology used was then revealed, with the clones containing more than one insert. Formation of concatamers can be avoided in future cDNA library construction by EcoRI methylation and restriction endonuclease treatment of the adapted cDNA prior to ligation into the cloning vector (Hodge *et al.*, 1992).

Sixty percent of the differentially expressed sequences, specific to both core and peripheral tissues of the mature sugarcane culm, remained unidentified, with no significant sequence similarity to any known plant genes. A high percentage of non–identified cDNA sequences was also observed in other similar projects. For example, 69.4% from rice endosperm (Liu *et al.,* 1995), 80% from maize leaf (Keith *et al.,* 1993) and 68% from *Arabidopsis thaliana* (Newman *et al.,* 1994). There are several possible reasons to explain the low number of identified cDNAs. Firstly, a stretch of high amino acid identity over a short sequence region can be more meaningful than a stretch of lower amino acid identity over the entire sequence region, if the short sequences represent a domain which determines biological function (eg. bZIP transcription factors; Hurst, 1994). Secondly, the localisation of the potential conserved domain(s) is important; if a domain of high sequence similarity is

not contained within the 5' region, but rather localised in the middle or in the C–terminal part of the gene (eg. HD –ZIP transcription factors; Meissner and Theres, 1995), partial sequencing of cDNA ends will not allow the identification of the corresponding gene. A further consideration is that plant genes are underrepresented in the databases compared with entries from other organisms.

Analysis of the ESTs of the co-expressed cDNAs with a relatively stronger pattern of gene expression in the core as compared to the peripheral tissues of the mature culm, encoded for genes with a wide range of cellular metabolism (protein synthesis, protein modification, structural protein) while the majority of the genes were found to be associated with cell wall metabolism, signal transduction and stress responses. Additionally, only one abundantly expressed transcript associated with sucrose metabolism (sucrose synthase) was identified. These results confirmed that culm maturation was not characterised by sucrose metabolism despite its distinct physiological characteristic of storing high levels of sugars. Furthermore, attempts to characterise the function of these genes in other plants, have not only confirmed their differential pattern of gene expression with respect to various stimuli (as briefly discussed below) but have provided an insight into the possible primary molecular mechanisms regulating gene expression in the mature culm.

For several cDNAs, significant sequence similarity was found to genes that had already been shown, in other plant systems, to be affected by abscisic acid (ABA), drought or other environmental stresses. Examples were genes such as the translation elongation factor (Rorat *et al.,* 1997), a water stress induced protein from cowpea (*Vigna unguiculata*) plants (Iuchi *et al.,* 2000) and a jacalin homolog, a lectin (Carson, 2001). Similarly, in maize, cellular stresses such as oxidative damage, osmotic stress or anaerobiosis occuring during cell wall digestion were found to have largely increased the transcripts encoding polyubiquitin genes (Garbarino *et al.*, 1992). Furthermore, identified cDNA homologous to a subfamily of protein kinases, the SNF1 homologs (sucrose non-fermenting 1) in plants were observed to play a role in the stress

response given their interactions with stress related calcium signals (Shi *et al.*, 1999). The regulatory role in cell metabolism was also observed in a developmental specific manner in response to metabolic signals such as nutrients (Shi *et al.*,1999). These were indicative of the possibility that differentially expressed genes in the mature culm, were simultaneously under developmental and stress regulation.

Furthermore, a number of cDNAs cloned in this study were homologous to genes expressed in response to water deficit such as water stressed and drought inducible proteins, which further affected the cell wall metabolism. Stress due to changes in osmotic potential may occur during low temperature treatments, pathogen infection, drought, salinity and due to storage of large amounts of osmotically active substances in storage tissues as in fruits (Davies and Robinson, 2000). For example in grape berry; water stressed and drought inducible proteins accumulated in response to stresses such as the storage of high concentrations of sugars in the vacuoles and rapid cell expansion or part of the ripening developmental program. This also paralleled the accumulation of hydroxyproline rich glycoprotein transcripts, a structural polymer involved in cell wall structure (Davies and Robinson, 2000). Furthermore, synthesis of another component of the cell wall metabolism such as structural polyprotein c identified in this study, known to be involved in providing additional support to the polysaccharide network in the cell wall by forming intermolecular cross links, had been reported to be upregulated in response to stress factors (Smart *et al.*, 1998). Besides strengthening the cell walls, they were produced in response to invasion attacks (Smart *et al.*, 1998). Thus, this suggested that some differentially expressed genes in mature culm tissues identified in this study were stress-responsive.

Furthermore, homologs of cDNAs identified in this study encode heat shock proteins (HSPs) and small heat shock smHSPs, which are a small group of evolutionary conserved polypeptides known to be synthesised in higher plants for protection against high temperature stresses and anaerobic responses. For example, Cooper and Ho (1987) found membrane associated smHSPs and HSP70 in maize and suggested that they have a potential role in

maintaining normal membrane processes during heat stress. On the other hand, some plant HSPs are required for normal growth at the upper end of their normal growth temperature range, whereas others help cells withstand the toxic effects of extreme temperatures (Sabehat *et al.*, 1998). However, there was also evidence that they are expressed with respect to specific developmental processes such as *tom66* and *tom11,* also described as ripening-related genes in tomato (Fray *et al.*, 1990). An example of such a gene isolated in this study was ABA ripening inducible gene.

Hence, these results suggested that the differential expression of highly abundant specific genes in the mature culm, might be regarded as an adaptive response to a wide variety of stress conditions. During culm maturation, cells rapidly expand since they have to accumulate large amounts of the sucrose and water. This consequently affects osmotic pressure and water potential, thus making the plant vulnerable to the reduction in water activity and an increase in temperature. Therefore, part of the adjustment to the rapid increase in vacuolar sugar levels may be the synthesis of proteins involved in stress management including those playing a role in internodal expansion, culm maturation and protection against environmental stimuli, which may dramatically affect their survival and productivity.

5.5 Differential Gene Expression in the Mature Culm using northern analysis

Although the cDNA expression arrays have emerged as a popular tool for gene expression studies, there are factors that can limit the amount of information obtainable through this approach (Southern *et al.*, 1999). To evaluate the validity of results obtained through array screening in an independent manner, clones S CLO 14, 16, 22, 23 and 25 containing 9 unique differentially expressed sequences were tested using northern hybridisation analysis. A distinct pattern of mRNA accumulation observed in the core and peripheral tissues confirmed the presence of specific regulatory elements in the promoter regions, controlling late developmental stage specific expression in the sugarcane. However, results obtained from the expression arrays did

not coincide with the northern hybridisation analysis results for all the cDNA fragments characterised. For example, the selective pattern of gene expression as observed on the arrays (that is cDNAs showing a stronger gene expression pattern in the core tissues as compared to the peripheral tissues) was not detected on the northern hybridisation analysis membranes, indicating some quantitative variation in the two cases. It remains to be seen whether these differences are the results of the hybridisation of the behaviour of single DNA fragments used for northern blotting and the complex mixtures of DNA fragments used for the DNA arrays. Hence, observations made in this study revealed that cDNA gene expression analysis was more of a qualitative rather than quantitative gene expression assay for the identification of differentially expressed genes in the mature culm.

Furthermore, tissue specific genes in sugarcane were identified using cDNA expression arrays. In contrast, expression of these genes were not detected on the northern analysis, indicating that gene transcription was down-regulated or turned off by specific regulatory components in the promoter region of the respective genes. The cDNA that was synthesised and used for cDNA library construction and the labelled cDNA representations for array screening were from the same mRNA populations derived from sugarcane tissues grown and harvested over summer seasons. However, for northern analysis, the RNA samples used were extracted during the winter season. Hence, the variation of expression patterns in the two cases could represent sequences that are actively transcribed in summer while their transcription is down-regulated or turned off in winter.

Plant response to temperature changes leads to alteration of the cellular metabolism and is correlated to significant changes in gene expression (for reviews see Guy, 1990; Palva, 1994). Cold-regulated genes have been isolated and characterised in several plant species including alfalfa (Wolfraim *et al.,* 1993), *Arabidopsis* (Hajela *et al.,* 1990), spinach (Neven *et al.,* 1993), tomato (Schaffner and Fischer, 1990), barley (Cattivelli and Bartels, 1990) and wheat (Houde *et al.,* 1992). In the *Arabidopsis rd29a* promoter, a dehydration-responsive element (DRE) was identified (TACCGACAT). Within

1 kb of the transcription site another two conserved sequence elements were identified and six G-box motifs (CACGTC) were found to be spread over the whole promoter (Kirch *et al.*, 1997). Thus, tissue and peripheral specific genes identified in this study could be used for the isolation of stress-inducible promoters (i.e.heat-induced) in sugarcane while selectively expressed genes identified using the cDNA analysis can be used for mature culm specific promoter isolation.

CONCLUDING REMARKS

Identification and characterisation of differentially expressed sequences from two specific tissue types of the mature sugarcane culm was undertaken in the present investigation. Since an altered pattern of morphological changes and sucrose accumulation was previously observed in the sugarcane mature culm, it was hypothesised that genes were differentially expressed.

The following were revealed;

1. A combination of differential screening, DNA sequence analysis and northern analysis allowed the rapid characterisation of a number of differentially expressed gene sequences that were specific to the core and the peripheral tissues of the mature sugarcane culm.

2. Culm maturation was characterised by genes expressed in response to a wide range of stresses. The diversity of differentially expressed genes identified in this study provided new insights into the genetic regulation of mature culm development. In particular, the lack of association between the types of genes identified as preferentially expressed in the mature culm and available biochemical and physiological data regarding sucrose accumulation in the sugarcane illustrated the complexity of sugarcane metabolism at gene level. In-depth molecular analyses are required to develop an understanding of how sucrose metabolism is regulated at the level of gene transcription and translation.

Future Directions

1. Large scale sequencing of the two developed specific cDNA libraries representative of the mature sugarcane culm can be performed to identify and characterise new sugarcane genes in order to increase the availability of plant genes in the international public database.

2. Identification of new sugarcane genes expressed specifically in the mature culm can be beneficial for sugarcane mapping and genetic manipulation.

3. Characterised genes can be used as research tools in understanding the complexity of internodal expansion and culm maturation.

4. Lastly, the long term goal of this study can be achieved by using the identified differentially expressed genes for the isolation of tissue-specific, stage-specific and stress-inducible promoters in the mature sugarcane culm.

REFERENCES

Aguan, K., Sugawara, K., Suzuki, N. and Kusano, T. 1991 Isolation of genes for low temperature-induced proteins in rice by a simple subtracted method. *Plant and Cell Physiology*, **32**: 1285 – 1289.

Albert H., Zhu Y.J., Carr J. and Moore P.H. 1996 Structure and expression of soluble acid invertase genes in the stem of high and low sucrose accumulating *Saccharum* species and hybrids. In '*Sugarcane: research towards efficient and sustainable production*'. (Eds JR Wilson, DM Hogarth, JA Campbell and AL Garside). 102 – 104. (CSIRO Division of Tropical Crops and Pastures: Brisbane)

Altschul, S.F., Gish, W., Miller, W., Meyrs, E.W. and Lipman, D. 1990 Basic local alignment search tool. *Journal in Molecular Biology,* **215**: 403 - 410.

An, G. and Kim, Y. 1993 Techniques for isolating and characterising plant transcription promoters, enhancers and terminators. In *Methods in plant molecular biology and biotechnology*. Glick, B. R. and Thompson, J. E. (Eds.), ISBN 0-8493-5164-2: 155 – 166. C.R.C. Press, London.

An, G., Luo, G., Veltri, R.W. and O'Hara, S.M. 1996 Sensitive, nonradioactive differential display method using chemiluminescent detection. *Biotechniques,* **20**: 342 – 346.

Andrews, D.L., Beames, B., Summers, M.D. and Park, W.D. 1988 Characterisation of the lipid acyl hydrolase activity of the major tuber protein patatin by cloning and expression in a baculovirus vector. *Biochem. J.,* **152**: 199 – 206.

Ausubel, 1996 Current Protocols in Molecular Biology 2, John Wiley and Sons, New York, Chapter 4.

Bachem, C.W., van der Hoeven, R.S., de Bruijn, S.M., Vreugdenhil, D., Zabeau, M. and Visser, R.G. 1996 Visualisation of differential gene expression using a novel method of RNA fingerprinting based on AFLP: analysis of gene expression during potato tuber development. *Plant Journal,* **9**: 745 – 753.

Bauer, D., Muller, H., Reich, J., Reidel, H., Ahrenkiel, V., Warthoe, P. and Strauss, M. 1993 Identification of differentially expressed mRNA species by an improved display technique (DDRT-PCR). *Nucleic Acids Research,* **21**: 4272 – 4280.

Benfey, P.N. and Chua, N.H. 1989 Regulated Genes in transgenic plants. *Science,* **244**: 174 – 181.

Benfey, P.N. and Chua, N.H. 1990 The Cauliflower mosaic virus 35S Promoter: Combinatorial regulation of transcription in plants. *Science,* **250**: 959 – 966.

Bower, R. and Birch, R.G. 1992 Transgenic sugarcane plants via microprojectile bombardment. *The Plant Journal,* **2**: 409 – 416.

Boutry, R. and Chua, N.H. 1985 A nuclear gene encoding the beta subunit of the mitochondrial ATPase in *N. plumbaginifolia. EMBO Journal,* **4**: 2159 – 2165.

Brown, P.O. and Botstein, D. 1999 Exploring the new world of the genome with DNA microarrays. *Nature Genetics,* **21**: 33 – 37.

Callis, J., Fromm, M. and Walbot, V. 1987 Intron increase gene expression in cultured maize cells. *Genes and Development,* **1**: 1183 – 1200.

Carson, D. and Botha, F.C. 2000 Preliminary analysis of expressed sequence tags for sugarcane. *Crop Science,* **40**: 1769 – 1779.

Carson, D, 2001 Gene discovery and expression analysis in sugarcane leaf and culm. (PhD. Thesis – University of Natal).

Cattivelli, L. and Bartels, D. 1990 Molecular cloning and characterisation of cold-regulated genes in barley. *Plant Physiology*, **93**: 1504 – 1510.

Chen, J.J., Wu, R., Yang, P.C., Huang, J.Y., Sher, Y.P., Han, M.H., Kao, W.C., Lee, P.I., Chiu, T.F., Chang, F., Chu, Y. W., Wu, C.W. and Peck, K. 1998 Profiling expression patterns and isolating differentially expressed genes by cDNA microarray system with colorimetry detection. *Genomics*, **51**: 313 – 324.

Chetverin, A.B. and Kramer, F.R. 1994 Oligonucleotide arrays: new concepts and possibilities. *Biotechnology (N. Y.)*, **12**: 1093 – 1099.

Christensen A.H., Sharrock R.A. and Quail P.H. 1992 Maize polyubiquitin genes: structure, thermal pertubation of expression and transcript splicing, and promoter activity following the transfer to protoplasts by electroporation. *Plant Molecular Biology*, **18**: 675 – 689.

Christopher P.L.G. and Campbell J.A., 2001 Sugarcane sucrose metabolism: scope for molecular manipulation. *Australian Journal of Plant Physiology*, **28**: 1 – 12.

Clark, A.M., Jacobsen, K.R., Bostwick, D.E., Dannenhoffer, J.M., Skaggs, M.I. and Thompson, G.A. 1997 Molecular characterisation of a phloem-specific gene encoding the filament protein. Phloem Protein 1 (1), from the *Curcurbita maxima*. *The Plant Journal*, **11**: 49 – 61.

Cooper, P. and Ho, T. 1987 Intracellular localization of heat shock proteins in maize. *Plant Physiology*, **84**: 1197 – 1203.

Cooke, R., Raynal, M., Laudie, M., Grellet, F., Delseny, M., Morris, P.C., Guerrier, D., Giraudat, J., Quigley, F., Clabault, G., Li, Y.F., Mache, R.,

Krivitzky, M., Kreis, M., Lecharny, A., Parmentier, Y., Marbach, J., Fleck, J., Clement B., Phillips, G., Herve, C., Bardet, C., Tremousay, D., Lescure, B., Lacomme, C., Roby, D, Jourgon, M.F., Chabrier, P., Charpenteau, J., Desprez, T., Aselem, J., Chiapello, H. and Hofte, H. 1996 Further progress towards a catalogue of all *Arabidopsis* genes: analysis of a set of 5000 non-redundant ESTs. *Plant Journal,* **9**: 101 – 124.

Cornejo, M., Luth, D., Blankenship, K.M., Anderson, O.D. and Blechl, A.E. 1993. Activity of a maize Ubiquitin promoter in transgenic rice. *Plant Molecular Biology,* **23**: 567 – 581.

Coupe, S.A. and Deikman, J. 1997 Characterisation of a DNA-binding protein that interacts with 5' flanking regions of two fruit-ripening genes. *The Plant Journal*, **11**: 1207 – 1218.

Croy, E.J., Ikemura, T., Shirsat, A. and Croy, R.R.D. 1993 Plant nucleic acids, in Croy R.R.D (Ed). *Plant Molecular Biology Labfax*, 21 – 48, Oxford : Bios Scientific Publishing Ltd.

Custers, J.B.M., Oldenhof, M.T., Schrauwen, J.A.M. and Cordewcner, J.H.G. 1997 Analysis of microspore-specific promoters in transgenic tobacco. *Plant Molecular Biology*, **35**: 689 – 699.

Dale, P.J. 1995 R & D regulation and field trialling of transgenic crops. *Tibtech,* **13**: 398 – 403.

Datson, N.A., van der Perk-de Jong, J., van den Berg, M.P., de Kloet, E.R. and Vreugdenhill, E. 1999 MicroSAGE: a modified procedure for serial analysis of gene expression in limited amounts of tissue. *Nucleic Acids Research*, **27**: 1300 – 1307.

Davies, C. and Robinson, S.P. 1996 Sugar accumulation in grape berries: cloning of two putative vacuolar invertase cDNAs and their expression in grapevine tissue. *Plant Physiology,* **111**: 275 – 283.

Davies, C. and Robinson, S.P. 2000 Differential screening indicates a dramatic change in mRNA profiles during grape berry ripening. Cloning and characterization of cDNAs encoding putative cell wall and stress response proteins. *Plant Physiology*, **122**: 803 – 812.

Decraene, C., Reguigne-Arnould, I., Auffray, C. and Pietu, G. 1999 Reverse transcription in the presence of dideoxynucleotides to increase the sensitivity of expression monitoring with cDNA arrays. *Biotechniques*, **27**: 962 – 966.

Dellaporta, S.L., Wood, J. and Hicks, J.B. 1983 A plant DNA minipreparation: Version II. *Plant Molecular Biology Reporter,* **1**: 19-21.

DeRisi, J.L., Lyer, V.R. and Brown, P.O. 1997 Exploring the metabolic and genetic control of gene expression on a genomic scale. *Science*, **278**: 680 – 686.

DeSaizieu, A., Certa, U., Warrington, J., Gray, C., Keck, W. and Mous, J. 1998 Bacterial transcript imaging by hybridisation of total RNA to oligonucleotide arrays. *Nature Biotechnology*, **16**: 45 – 48.

Desprez, T., Amselem, J., Caboche, M. and Hofte, H. 1998 Differential gene expression in *Arabidopsis* monitored using cDNA arrays. *The Plant Journal,* **14**: 643 – 652.

Donohue, P.J., Alberts, G.F., Guo, Y. and Winkles, J.A. 1995 Identification by targeted differential display of an immediate early gene encoding a putative serine/threonine kinase. *Journal of Biological Chemistry*, **270**: 10351 – 10357.

Duggan, D.J., Bittner, M., Chen, Y., Meltzer, P. and Trent, J.M. 1999 Expression profiling using cDNA microarrays. *Nature geneticst,* **21**: 10 – 14.

Eisen, M.B., Spellman, P.T., Brown, P.O. and Botstein, D. 1998 Cluster analysis and display of genome-wide expression patterns. *Proceedings of the National Academy of Science, USA*, **95**: 14863 – 14868.

Elliott, A.R., Campbell, J.A., Brettell, R.I. and Grof, C.P.L. 1998 *Agrobacterium*-mediated transformation of sugarcane using GFP as a screenable marker. *Australian Journal of Plant Physiology,* **25**: 739 – 743.

Fang, R.X., Nagy, F., Sivasubramanian, S. and Chua, N.H. 1989 Multiple *cis* regulatory elements for maximal expression of the cauliflower mosaic virus 35S promoter in transgenic plants. *Plant Cell,* **1**: 141 – 150.

Fillion L., Ageorges A., Picaud S., Coutos – Thevenot P., Lemoine R., Romieu C. and Delrot S. 1999 Cloning and expression of a hexose transporter gene expressed during the ripening of grape berry. *Plant Physiology*, **120**: 1083 – 1093.

Finnegan J. and McElroy D. 1994 Transgene Inactivation: Plants fight back! *Bio/Technology*, **12**: 883 – 888.

Fischer, A., Saedler, H. and Thiessen, G. 1995 Restriction fragment length polymorphism-coupled domain-directed differential display: a highly efficient technique for expression analysis of multigene families. *Proceedings of the National Academy of Sciences of the United States of America* **92**: 5331 – 5335.

Foder, S. P., Read, J. L., Pirrung, M. C., Stryer, L., Lu, A. T. and Solas, D. 1991 Light-directed, spatially addressable parallel chemical synthesis. *Science*, **251**: 767 – 773.

Fray, R.G., Lyeet, G.W. and Grierson, D. 1990 Nucleotide sequence of a heat shock and a ripening related cDNA from tomato. *Nucl Acids Res*, **18**: 7148 – 7153.

Freeling, M. and Bennet, D.C. 1985 Maize *Adh*1. *Annual Review in Genetics,* **19**: 297 – 323.

Fromm, M. F., Morrish, F., Armstrong, C., Williams, R., Thomas, J. and Klein, T.M. 1990 Inheritance and expression of chimeric genes in the progeny of transgenic maize plants. *Bio/Technology,* **8**: 833 – 839.

Gallo-Meagher, M., and Irvine, J.E. 1993 Effects of tissue type and promoter strength on transient GUS expression in sugarcane following particle bombardment. *Plant Cell Reports*, **12**: 666 – 670.

Galun, E. and Breiman, A. 1997 Transgenic Plants; with an appendix on intellectual properties & commercialisation of transgenic plants by John Barton. Imperial College Press. London, Chapter 4.

Garbarino, J.E., Rockhold, D.R. and Belknap, W.R. 1992 Expression of stress-responsive ubiquitin genes in potato tubers. *Plant Molecular Biology,* **20**: 235 – 244.

George, K.S., Zhao, X., Gallahan, D., Shirkey, A., Zareh, A. and Esmaeli-Azad, B. 1997 Capillary electrophoresis methodology for identification of cancer related gene expression patterns of fluorescent differential display polymerase chain reaction. *Journal of Chromatography B*, **695**: 93 – 102.

Gittins, J.R., Pellny, T.K., Hiles, E.R., Rosa, C., Biricolti, S. and James, D.J. 2000 Transgene expression driven by heterologous ribulose–1,5–bisphosphate carboxylase / oxygenase small – subunit gene promoters in the vegetative tissues of apple (*Malus pumila* Mill). *Planta*, **210**: 232 – 240.

Giuliano, G., Pichersky, E., Malik, V.S., Timko, M.P., Scolnik, P.A. and Cashmore, A.R. 1988 An evolutionary conserved protein binding sequence upstream of a plant light-regulated gene. *Proc. Natl. Acad. Sci. USA,* **85**: 7089 – 7093.

Gonsky, R., Knauf, J.A., Elisei, R., Wang, J.W., Su, S. and Fagin J.A. 1997 Identification of rapid turnover transcripts over-expressed in thyroid tumors and thyroid cancer cell lines: use of a targeted differential RNA display

method to select for mRNA subsets. *Nucleic Acids Research*, **25**: 3823 – 3831.

Gordon – Kamm W.J., Micahel Spencer, T., Mangano, M.L., Adams, T.R., Daines, R.J., Start, W.G., O'Brien, J.V., Chambers, S.A., Adams, Jr.W.R., Willets, N.G., Rice, T.B., Mackey, C.J., Krueger, R.W., Kausch, A.P. and Lemaux, P.G. 1990 Transformation of maize cells and regeneration of fertile transgenic plants. *The Plant Cell*, **2**: 603 – 618.

Gottschlich, S., Goeroegh, T., Folz, B.J., Lippert, B.M. and Werner, J.A. 1997 Optimised differential display and reamplification parameters for silver staining. *Research Communications in Molecular Pathology and Pharmacology*, **97**: 237 – 240.

Grof, C.P.L., Campbell, J.A., Quick, W.P., Sonnewald, U. and Glassop, D., 1996a Molecular manipulation of sucrose phosphate synthase in sugarcane. In '*Sugarcane: research towards efficient and sustainable production*'. (Eds J.R. Wilson, D.M. Hogarth, J.A. Campbell and A.L. Garside) 124 – 126. (CSIRO Division of Tropical Crops and Pastures: Brisbane)

Guy, C.L. 1990 Cold acclimation and freezing stress tolerance: role of protein metabolism. *Annual review of Plant Physiology and Plant Molecular Biology,* **41**: 187 – 223.

Hajela, R.K., Horvath, D.P., Gilma, S.J. and Thomashow, M.F. 1990 Molecular Cloning and expression of *cor* (Cold-regulated) genes in *Arabidopsis thaliana. Plant Physiology*: **125**: 1246 – 1252.

Hanler, B.A. and Schuler, M.A., 1988 Plant intron sequences: evidence for distinct groups of introns. *Nucl. Acids Res.* **16**: 7159 – 7176.

Hansom, S., Bower, R., Zhang, L., Potier, B., Elliot, A., Basnayake, S., Cordeiro, G., Hogarth, D.M., Cox, M., Berding, N. and Birch, R.G. 2001 Regulation of transgene expression in sugarcane. *BMb,* **5**: 278 – 289.

Heineke, D., Sonnewald, U., Bussis, D., Gunter, G., Leidreiter, K., Wilke, I., Raschke, K., Willmitzer, L. and Heldt, H.W. 1992 Apoplastic expression of yeast derived invertase in potato. Effects on photosynthesis, leaf solute composition, water relations and tuber composition. *Plant Physiology,* **100**: 301 – 308.

Hodge, R., Paul, W., Draper, I. and Scott, R. 1992 Cold plaque screening: a simple technique for the isolation of low abundance, differentially expressed transcripts from conventional cDNA libraries. *The Plant Journal*, **2**: 257 – 260.

Houde, M., Danyluk, J., Rassart, E., Dhindsa, R.S. and Sarhan, F. 1992 Cloning, characterisation and expression of a cDNA encoding a 50-kilodalton protein specifically induced by cold acclimation in wheat. *Plant Physiology*, **99**: 1381 – 1387.

Hurst, H.C. 1994 bZIP proteins. *Protein Profile*, **1**: 123 – 168.

Iuchi, S., Kobayashi, M., Yamaguchi-Shinozaki, K. and Shinozaki, K. 2000 A stress-inducible gene for 9-cis-epoxycarotenoid dioxygenase involved in abscisic acid biosynthesis under water stress in drought-tolerant cowpea. *Plant Physiology*, **123**: 553 – 562.

Jacobsen, K.R., Fisher, D.J., Maretzki, A. and Moore, P.H. 1992 Developmental changes in the anatomy of sugarcane stem in relation to phloem unloading and sucrose storage. *Botanica Acta,* **105**: 70 – 79.

Jefferson, R., Goldsborough, A. and Bevant, M. 1990 Transcriptional regulation of a patatin-1 gene in potato. *Plant Molecular Biology*, **14**: 995 – 1006.

Jepson, I., Bray, J., Jenkins, G., Schuch, W. and Edwards, K. 1991 A rapid procedure for the construction of PCR cDNA libraries from small amounts of plant tissue. *Plant Molecular Biology Reporter,* **9**: 131 – 138.

Keith, B. and Chua, N.H. 1986. Monocot and dicot pre-mRNAs are processed with different efficiencies in transgenic tobacco. *EMBO,* **J5**: 2149 – 2425.

Keith, C.S., Hoang, D.O., Barrett, B.M., Feigelman, B., Nelson, H., Thai, H. and Baysdorfer, C. 1993 Partial sequence analysis of 130 randomly selected maize cDNA clone. *Plant Physiology* **103**: 329 – 332.

Kirch, H.H., Van Berkel, J., Glaczinski, H., Salamini, F. and Gebhardt, C. 1997 Structural organisation, expression and promoter activity of a cold-stress-inducible gene of potato (*Solanum tuberosum L.*). *Plant Molecular Biology*, **33**: 897 – 909.

Kuhlemeier, C., Green, P. and Chua, N.H. 1987 Regulation of gene expression in higher plants. *Annual Review of Plant Physiology*, **38**: 221 – 257.

Kuhn, E. 2001 From library screening to microarray technology: Strategies to determine gene expression profiles and to identify differentially regulated genes in plants. *Annals of Botany*, **87**: 139 – 155.

Kyozuka, J., Izawa, T., Nakajima, M. and Shimamoto, K. 1990 Effect of the promoter and the first intron of maize *Adh1* on foreign gene expression in rice. *Maydica,* **35**: 353 – 357.

Last, D.I., Brettell, R.I.S., Chamberlain, D.A., Chaudhury, A.M., Larkin, P.J., Marsh, E.L., Peacock, W.J. and Dennis, E.S. 1991 PEmu: An improved promoter for gene expression in cereal cells. *Theor. Al. Genet,* **81**: 581 – 588.

Ledoux, L. and Huart, R. 1969 Fate of exogenous bacterial deoxynucleic acids in barley seedlings. *J. Mol. Biol,* **43**: 243 – 262.

Lefebvre, D.D. and Gellatly, K.S. 1997 Fundamentals of gene structure and control. In Plant Metabolism (2nd. Ed.) Dennis. D.T., Turpin, D.H., Lefebvre,

D.D. and Layzell, D.B. (Eds.), ISBN 0-582-25906-1, 3 – 16. Longman Singapore Publishers, Singapore.

Liang, P. and Pardee, A.B. 1992 Differential display of eukaryotic messenger RNA by means of the polymerase chain reaction. *Science*, **257**: 967 – 971.

Liu, J., Hara, C., Umeda, M., Zhao, Y. and Okita, T.W. 1995 Analysis of randomly isolated cDNAs from developing endosperm of rice (*Oryza sativa* L.): evaluation of expression sequence tags and expression levels of mRNAs. *Plant Molecular Biology Reporter,* **2**: 685 – 689.

Liu, Z.B., Ulmasov, T., Shi, X., Hagen, G. and Guilfoyle, T.J. 1994 Soybean GH3 promoter contains multiple auxin-inducible elements. *Plant Cell* **6**: 645 – 657.

Lockhart, D.J. *et al.* 1996 Expression monitoring by hybridisation to high-density oligonucleotide arrays. *Nature Biotechnology,* **14**: 1675 – 1680.

Lockhart, D.J. and Winzeler, E.A. 2000 Genomics, gene expression and DNA arrays. *Nature*, **405**: 827 – 836.

Lohmer, S., Maddaloni, M., Motto, M., Di Fonzo, N., Hartings, H., Salamini, F. and Thompson, R.D. 1991 The maize regulatory locus *opaque-2* encodes a DNA-binding protein which activates the transcription of the *b-32* gene. *EMBO*, **10**: 617 – 624.

Marcotte, W.R., Russell, S.H. and Quantrano, R.S. 1989 Abscisic acid responsive sequence from the *Em* gene of wheat. *Plant Cell,* **1**: 969 – 976.

Martin-Laurent, F., Van Tuinen, D., Dumas-Gaudot, E., Gianinazzi-Pearson, V., Gianinazzi, S. and Franken, P. 1997 Differential display analysis of RNA accumulation in arbuscular mycorrhiza of pea and isolation of a novel symbiosis-related plant gene. *Mol. Gen. Genet.*, **256**: 37 – 44.

Matzke, A.J.M., Neuhuber, F., Park, Y.D., Ambros, P.F. and Matzke, M.A. 1994 Homology dependent gene silencing in transgenic plants: Epistatic silencing loci contain multiple copies of methylated transgenes. *Mol. Gen. Genet.*, **244**: 219 - 229.

McCormick, A., Brady, H., Furkushima, J. and Karin, M. 1994 TATA binding protein associated factor 5 in TFII D function through the initiator to direct basal transcription from a TATA less class II promoter. *EMBO J.*, **13**: 3115 – 3126.

McNew, J.A. and Goodman, J.M. 1996 The targeting and assembly of peroxisomal proteins; some old rules do not apply. *Trends Biochem. Sci.*, **21**: 54 – 58.

Meissner, R. and Theres, K. 1995 Isolation and characterisation of the tomato homeobox gene THOM 1. *Planta,* **195**: 541 – 547.

Mittelsten Scheid, O., Jakovleva, L., Atsar, K., Malusynska, J. and Paszkowski, K. 1996 A change of ploidy can modify epigenetic silencing. *Proceedings of the National Academy of Science, U.S.A.,* **93**: 7114 – 7119.

Mogen, B.D., MacDonald, M.H.J., Leggewie, G. and Hunt, A.G. 1992 Several distinct types of sequence elements are required for efficient mRNA 3' end formation in a pea *rbc*S gene. *Molecular Cell Biology*, **12**: 5406 – 5414.

Morelli, G., Nagy, F., Fraley, R.T., Rogers, S.G. and Chua, N.H. 1985 A short conserved sequence is involved in the light-inducibility of a gene encoding ribulose 1,5 bisphosphate carboxylase small subunit of pea. *Nature*, **315**: 300 – 304.

Moore, P.H. 1995 Temporal and spatial regulation of sucrose accumulation in the sugarcane stem. *Australian Journal of Plant Physiology*, **22**: 661 – 671.

Neven, L.G., Haskell, D.W., Hofig, A., Li, Q.B., and Guy, C.L. 1993 Characterisation of a spinach gene responsive to low temperatures and water stress. *Plant Molecular Biology*, **21**: 291 – 305.

Newman, T.E., De Bruijn, F.J., Green, P., Keegstra, K., McIntosh, L., Ohlrogge, J., Raikhel, S., Somerville, S., Thomashow, N., Retzel, E. and Somerville, C. 1994 Summary of methods for accessing results from large-scale sequencing of anonymous *Arabidopsis* cDNA clones. *Plant Molecular Biology,* **106**: 1241 – 1255.

Nover, L. 1994 Plant Promoters and transcription factors. Springer-Verlag, Heidelberg.

Odell, J.T., Nagy, F. and Chua, N.H. 1985 Identification of DNA sequences required for activity of the cauliflower mosaic virus *35S* promoter. *Nature*, **313**: 810 - 812.

Omirulleh,S., Abraham, M., Golovkin, M., Stefanov, I., Karabaev, M.K., Mustardy, L., Morocz, S. and Dudits, D. 1993 Activity of a chimeric promoter with the doubled CaMV 35S enhancer element in protoplast-derived cells and transgenic plants in maize. *Plant Molecular Biology*, **21**: 415 – 428.

Palva, E.P. 1994 Gene expression under low temperature stress. In stress-induced gene expression in plants. Basra, A. S. (Ed), ISB 3-7186-5466-0, pp 103 – 130. Harwood Academic publishers, Switzerland.

Pease, A.C. *et al.* 1994 Light-generated oligonucleotide arrays for rapid DNA sequence analysis. *Proceedings of the National Academy of Science, U.S.A.,* **91**: 5022 – 5026.

Pla. M., Vilardell, J., Guiltinan, M.J., Marcotte, W.R. and Niogret, M.F. 1993 The *cis* regulatory element CCACGTGG is involved in ABA and water-stress responses of the maize gene *rab28*. *Plant Mol. Biol.*, **21**: 259 – 266.

Rathus, C., Bower, R. and Birch, R.G. 1993 Effects of promoter, intron and enhancer elements on transient gene expression in sugarcane and carrot protoplasts. *Plant Molecular Biology*, **23**: 613 – 618.

Rorat, T., Irzykowski, W. and Grygorowicz, W.J. 1997 Identification and expression of novel cold induced genes in potato. *Plant Science*, **124**: 69 – 78.

Rosenthal, N. 1987 Identification of regulatory elements of cloned genes and functional assays. *Methods in Enzymology*, **152**: 704 – 720.

Ruan, Y., Gilmore, J. and Conner, T. 1998 Towards *Arabidopsis* genome analysis: monitoring expression profiles of 1400 genes using cDNA microarrays. *Plant Journal*, **15**: 821 – 833.

Sambrook, J., Fritsch, E.F. and Maniatis, T. 1989 Molecular Cloning: A laboratory manual, 1989. Cold Spring Harbour Laboratory. Cold Spring Harbour, NY.

Sabehat, A., Lurie, S. and David, W. 1998 Isolation and characterization of a heat induced gene, *hcit2,* encoding a novel 16.5 kDa protein: expression coincides with heat induced tolerance to chilling stress. *Plant Molecular Biology*, **36**: 935 – 939.

Salgueiro, S., Pignocchi, C. and Parry, M.A.J. 2000 Intron-mediated GUS expression in tritordium and wheat resulting from particle bombardment. *Plant Molecular Biology,* **42**: 615-622.

Sarni-Manchado, P., Verries, C. and Tesniere, C. 1997 Molecular characterization and structural analysis of one alcohol dehydrogenase gene (GV-*Adh1*) expressed during ripening of grapevine (*Vitis vinifera* L.) berry. *Plant Science,* **125**: 177-187.

Sasaki, T., Song, J., Koga-Ban, Y., Matsui, E., Fang, F., Higo, H., Nagasaki, H., Hori, M., Miya, M., Murayama-Kayano, E., Takiguchi, T., Takasuga, A., Niki, T., Ishimaru, K., Ikeda, H., Yamamoto, Y., Mukai, Y., Ohta, I., Miyadera, N., Havukkula, I. and Minobe, Y. 1994 Toward cataloguing all rice genes: large scale sequencing of randomly chosen rice cDNAs from a callus cDNA library. *Plant Journal*, **6**: 615 – 624

Schaffner, M.A., and Fischer, R.L. 1990 Transcriptional activation by heat and cold of a thiol protease gene in tomato. *Plant Physiology*, **93**: 1486 – 1491.

Schena, M., Shalon, D., Davis, R. and Brown, P. 1995 Quantitative monitoring of gene expression patterns with a complimentary DNA microarray. *Science*, **270**: 467 – 470.

Schulze-Lefert, P., Becker-Andre, Schulz, M., Hahlbrock, K. and Dangl, J.L. 1989 Functional architecture of the light responsive chalcone synthase promoter from parsley. *Plant Cell* **1**: 707 – 714.

Shen, Q. and Ho, T.D. 1995 Functional dissection of an abscisic acid (ABA) inducible gene reveals two independent ABA-responsive complexes each containing a G-box and a novel *cis*-acting element. *Plant Cell*, **7**: 295 – 307.

Shi, J., Kim, K., Ritz, O., Albrecht, V., Gupta, R., Harter, K., Luan, S. and Kudla, J. 1999 Novel protein kinases associated with calcineurin B-like calcium sensors in Arabidopsis. *The Plant Cell*, **11**: 2393 – 2405.

Shimamoto, K., Terada, R., Izawa, T. and Fujimoto, H. 1989 Fertile transgenic rice plants regenerated from transformed protoplasts. *Nature,* **338**: 274 – 276.

Simpson, C.G., Clark, G., Davidson, D., Smith, P. and Brown, J.W.S. 1996 Mutation of putative branch points consensus sequences in plant intron reduces splicing efficiency. *Plant Journal,* **9**: 369 – 380.

Smale, S.T. and Baltimore, D. 1989 The "initiator" as transcription control element. *Cell*, **57**: 103 – 113.

Smart, L.B., Vojdani, F., Maeshima, M. and Wilkins, T.A. 1998 Genes involved in osmoregulation during turgor-driven cell expansion of developing cotton fibres are differentially regulated. *Plant Physiology,* **116**: 1539 – 1549.

Sonnewald, U. and Willmitzer, L. 1992 Molecular approaches to sink-source molecular interactions. *Plant Physiology*, **99**: 1267 – 1270.

Southern, E., Mir, K. and Shchepinov, M. 1999 Molecular interactions on microarrays. *Nature Genet.*, **21**: 5 – 9.

Sullivan, M.L. and Green, P.J. 1993 Post-transcriptional regulation of nuclear encoded genes in higher plants: The roles of mRNA stability and translation. *Plant Molecular Biology,* **23**: 1091 – 1104.

Tavazoie, S., Hughes, J.D., Campbell, M.J., Cho, R.J. and Church, G.M. 1999 Systematic determination of genetic network architecture. *Nature Genet.*, **22**: 281 – 285.

Tesniere, C. and Vayda, M.E. 1991 Method for the isolation of high quality RNA from grape berry tissues without contaminating tannins or carbohydrates. *Plant Molecular Biology,* **9**: 242 – 251.

Tjian, R. and Maniatis, T. 1994 Transcription activation. A complex puzzle with a few easy pieces. *Cell*, **79**: 5 – 8.

Van Dillewijn, C. 1952 Growth: general, grand period, growth formulae. In Botany of Sugarcane. Vol. 1. Van Dillewijn, C. (Ed.) 97 – 162. Veenen and Zonen, The Netherlands.

Vasil, I. K. 1994 Molecular improvement of cereals. *Plant Molecular Biology*, **25**: 925 – 937.

Vasil, V., Castillo, A.M., Fromm, M.E., and Vasil, I.K. 1992 Herbicide resistant fertile transgenic wheat plants obtained by microprojectile bombardment of regenerable embryogenic callus. *Bio/Technology*, **10**: 667 – 674.

Velculescu, V.E., Zhang, L., Vogelstein, B. and Kinzler, K.W. 1995 Serial analysis of gene expression. *Science*, **270**: 484 – 487.

Vos, P., Hogers, R., Bleeker, M., Reijans, M., van de Lee, T., Hornes, M., Frijters, A., Pot, J., Peleman, J. and Kuiper, M. 1995 AFLP: a new technique for DNA fingerprinting. *Nucleic Acid Research,* **23**: 4407 – 4414.

Wan, Y. and Lemaux, P.G. 1994 Generation of large numbers of independently transformed fertile barley plants. *Plant Physiology,* **104**: 37 – 48.

Wan, J.S., Sharp, S.J., Poirier, G.M., Wagaman, P.C., Chambers, J., Pyati, J., Hom, Y.L., Galindo, J.E., Huvar, A., Peterson, P.A., Jackson, M.R. and Erlander, M.G. 1996 Cloning differentially expressed mRNAs. *Nature Biotechnology*, **14**: 1685-1691.

Wang, K., Gan, L., Jeffery, E., Gayle, M., Gown, A.M., Skelly, M., Nelson, P. S., Ng, W.V., Schummer, M., Hood, L and Mulligan, J. 1999 Monitoring gene expression profile changes in ovarian carcinomas using cDNA micrarray. *Gene*, **229**: 101 – 108.

Wenzler, H.C., Mignet, G.A., Fisher, L. and Park, W.D. 1989 Analysis of a chimeric class I patatin GUS gene in transgenic plants high level expression in tubers and sucrose inducible expression in a cultured leaf and stem explants. *Plant Molecular Biology,* **12**: 41 – 50.

Whittaker, A. and Botha, F.C. 1997 Carbon partitionjng during sucrose accumulation in sugarcane internodal tissue. *Plant Physiology,* **115**: 1651 – 1659.

Wodicka, L., Doug, H., Mittman, M., Ho, M.H., Lockhart, D.J. 1997 Genome wide expression monitoring in *Saccharomyces cerevisae*. *Nature Biotechnology*, **15**: 1354 – 1367.

Wolfraim, L.A., Langis, R., Tyson, H. and Dhindsa, R.S. 1993 cDNA sequence expression and transcript stability of a cold acclimation-specific gene, *cas 18*, of Alfalfa (*Medicago falcata*) cells. *Plant Physiology*, **101**: 1275 – 1282.

Yang,N.S and Russel, D. 1990 Maize sucrose synthase–1 promoter directs phloem cell specific expression of GUS gene in transgenic tobacco plants. *Proceedings of the National Academy of Sciences, U.S.A.*, **87**: 4144 – 4148.

Yershov, G., Barsky, V., Belgovskiy, A., Kirillov, E., Kreindlin, E., Ivanov, I., Parinov, S., Guschin, D., Drobishev, A., Dubiley, S., Mirzabekoy, A. 1996 DNA analysis and diagnostics on oligonucleotide microchips. *Proceedings of the National Academy of Sciences, U.S.A.*, **93**: 4913 – 4918.

Zhang, W., McElroy, D. and Wu, R. 1991 Analysis of rice Act1 5' region activity in transgenic rice plants. *Plant Cell,* **3**: 1155 – 1165.

Zhu, Q., Doerner, P.W. and Lamb, C.J. 1993 Stress induction and developmental regulation of a rice chitinase promoter in transgenic tobacco. *Plant Journal,* **3**: 203 – 212.

Zubay, G., 1993 DNA manipulation and its applications. In Biochemistry 3Rd Edition. Wm. C. Brown Publishers, Chapter 27.

APPENDIX

Nucleotide Sequences of the Putatively Identified Differentially Expressed Genes Specific to the Mature Sugarcane culm (sequences represented by their respective Clone Identity see Table 4.5).

ConSQ_1

AATTCCGATTTAGCCTCATAATCTTCATCTCCTTCGAATTCCGATTTAGCCTCATACTTCC
TATATCCTACGATGGGTGATGACTAGGACCAAGGGCATGGTCACCTCATATGTCTTACCT
GATGCTTTTGTTGGACATGACTACTATTTGGCACTGTAATCCAAACTAAGTGTGTGGCCT
CGAAACTATTTGCCTATGAGGCTAAATCGGAATTCCGAKKTAGCCTCATATTTTTTTTTTT
AAACCAACCACTTTATTTATTACTGATAATACTCGCAGAAATATGTACATGTACCGTGCAG
GCCATAAGGCACGGGCAACGATCGACGACCTGCAACTCCGAANACCGTTAGGAGTANTT
AWTCTTNACMCATTTATTTATTATTAGGTACAGACGGAACAGAAAGTTACGGGTTAGTCT
GGGCGTCCATCACGAACATGGTTCCGGTGRAAAGTTTGCATGCCATTAAGCTCCCGTAG
NANGTCCACACTGGTCGATCAAGCGTCGAAGTCCAAGCGCCACAGGTTAATCCCACAGT
TGGTAGARGCGCCCAGGGGGGGTATAGAAGGCTTGTATGTTGATCTTATACATCCTCCCG
CCATTAGTTTGGCGSACAACCTTGTGGTTCACAACACCGCGTGCCAACGTAAGTTGTCCA
ATTCCTCCAAAAATGCTCCACTGACCATCTTGTGGGATGTAACCCATCATCTGAAGCGTG
GATCCCTTCAGGCTCGTGCGCTCGAACACTATGGTCAAGGTGGTATACCAGGCTTGGTC
GTTCCTCTCGCTTGGNTTGAACTGNATGCCCTGCATACATCGTGCAACGACGANTRCGT
NGCTACCTTCCCCGATTTCCTCACGTACAATCCTGGCGGRTGTGTCCGAACCCSTCCCT
TGTTAGTGCTGAGTTCAGCCATTG

SCLO 23_a

CTCATCCAATTGGGTKTAACTCCATTATGTCACATGCCATTGAAACAACAGCGTACGCAG
TAAAGATCATAGATACAATTGGATGAGCTCATCACTAACAACAGACGGCGTTAGAACACC
AAGATCGGTCAGGAGAAGTGTGAGATATTGGGGAGGAGYGTAGTCTCTTGCAGATGTCT
MTCAGTCTTGATCTCATGTTATTGCCCAACATGAATGAATCTCCTGTTAATCTTCATCTGG
TACTCCCCTAGTGGATGCGAGTCGTTCGTAGAATTAAGCAAGA

SCLO 22_c

CATAAAAGGAAGGAATATTAAGGACTGCTTGGGACGGGCCTTTGAGTATCTCCATCAATG
TCATCATTCAAGGAGAGAGATCGTTATCCTTAAGTTAGATTTTGAAGAGGCCTTTGACCT
GGTGGAGCYATGAGGCTAAATCGCKATTCCGRTTGACCSTSATACACTCNCCCATTGRA
GAGAATTATGASTRCTTCCAAAGCATTKGAGTTGKTGCACACGGATCTATTTGGGCCAAC
TACATACACAAGCATTGATGGCAACAAGTATGGTTTTGTGATTGTGGATGATTACACTAG
ATACACATGGGTMTTCYTTCTTGNGGACCTATGAGGCTNAATSGGRACTCASATACATAC
AMSGACTGNTTGATAGCACCMRTCMCGGYCCAAGCAATTTATATAANCTATCAGCTCTAT
GAGGCTAAATCGGAATTTCGATTTAGCCTCATAGTAAGGCGATTCATNCATAATATTTATT
ATGTGCGGNTTCAAGATGCTCAAAAGATATATTTTGTAATCGGCTTTTGAGCTCCATTATG
TGTGCTTCATCAATTTAGGAATTATGGACTTGGCGGAGGGGGCAATACTAACTGCACCGT
TTAATGCTTCCTG

SCLO 23_c

CGCTTGCAGATCTTATTTATTATTTATTTAGAAGGGCAGTTGCATATTTCAATACCGCGTC
TCTCGTGTCATGCTGGGGATGACAACTTGGCTTAAGTGGCATGAGGGATAGGTTCGGCA
TTTTTGGCGCCGTTATCAGAATTAGAAAACTAAGTCTACTTTTGGTAATGACGTTAAGAAT
GCCCA

SCLO 22_d

CTTAAGACTGCACTAGATAATATTGATCTGCAGACGACCATTCTTTCTAGATTCGTTCTAA
TCTTTATCGTCAAAGATATCAGAATGTATGATCAAGATAAGCGAATAGCAAGCCAATTCTC
TCCATCACTTGCC

SCLO 23_d

CATAAAAGGAAGGAATATTAAGGACTGCTTGGGACGGGCCTTTGAGTATCTCCATCAATG
TCATCATTCAAGGAGAGAGATCGTTATCCTTAAGTTAGATTTTGAAGAGGCCTTTGACCT
GGTGGAGCC

SCLO 23_e

GAAGCTGGCGAGACTGAGGGATTATGGAAGCTGCTCAAGTGTCGCACAGAGGACGTGG
TGAAAATGCTGGTTTATCTCTGACGGTTCCTGAGGTGAACTCAATGGGTCTAGCATTATT
CGGAGGTTTAAGCTTAGAGC

SCLO 23_f

TGCTGGAACAGGTGGAGATGGAGGTTCTGACCACTGTCCGTAAGGCTCGGTTCCGCTGT
TGCGTCGCCCAACCGATGACGCTAGACTGAATCATCCACAATCCAGGAAAAGAACCTGG
AGAGAGGCTTAATCCTATGTATTTATAATATTATTAATAGACTTGTGCTTCAGAAGCAGAA
AGCTGACTCCTCAGTCTGCCTTCTGTCTTTGACATGTTATCGTAGAGTACACCTAGTTGT
CGTCACCAGAGTGGTTGTAATAGAGGTAGCTTTATGCTCGACTAAATACCCTTCCGTTGT
AA

SCLO 1_a

CGCGCACCTCTCCGCCGAGGACGAGGCGACGCTGGAGGCCGAGGTGCGGGGGTACTA
CGACGAGGCGGCGCCCAAGCGCCACACCAAGCCCTCCCGCAGCGAGCACTCCGCCGT
GTACGCCGACGCGCTCGTCCCGGACGCCGGCGGCAACACCCATCCGGAGCTCGACAA
GTTCCAAGAGCTCGAAGCACACACCGAGAGGTTGGCGTACGAGAGCGGCAATGTCGGA
GAGGAGTCCGTGGAGACGGAGTACTACAAGGATCTCGGCGGCGTTGGCAAGCAGCACC
ACACGACCGGAACGGGCTTCATCAAGATTGACAAAGCTAAGGGCGCCTCGTTCTAATTG
TCTGAAGACCCCAGCGCAGAGGAGCG

SCLO 1_b

GGTTGTTGCAAAGCAGGAAGGAGATAGCCAAGTATCTCTACGTGCCTCTCTTCAGCCTTT
GACAAGCCAGTTAGTCCATCCTTGACTTTGGGGATGATGAGGATGTGCGTTGGAGCTTG
AGGGTTTACATCCCTGAAAGCCAGTACCTTCTCGTCCTCATAAGCCACAGTAGAAGGGAT
TTCCTTTTTGATGATCTTGTCAAAT

SCLO 2_a

GAAAAACCGAACTGAAACTAATATAGAAACCCAACATAAGGGCAAGGTTGTGGCATGTCC
ACAAATACACCAAATAGAGAGGAAGAAAAACTACAACGCCAAGGCATATTAACTCTTCGA
TGATGCCTCCAAGGAGGTAAATGACGCCACTACGTAATCATCGTCGGTCAAGCAACAAG
TGGTGACGCTGAGCGCTTGGCAAGGCTTTCACCCGAGCCTTGTGTCGAGGGGTTGAAA
CAGGTCGACCATGGAGAAGAAGCAGCACAAA

SCLO 2_b

CGACTCTTCACCTTGTGCTCAGGCTCAGGGGTGGCATGCAAATCTTTGTCAAGACCCTC
ACTGGCAAGACCATCACCTTGGAGGTGGAGTCCTCGGACACCATTGACAATGTGAGGGC
GAAGATCCAGGACGAGGAGGGCATTCCCCCGGACCAGCAGCGTCTCATCTTCGCCGGC
AAGCAGCTTGAGGATGGCCGCACCCTTGCAGACTACAACATCCAGAAGGAGTCCACTCT
CCACCTGGTGCTCCGTCTCCSYGKKGKSSWGWATRSAKMYYKKSRYATGAGCAGCTGT
CCTTCCAGGTTCACAAGTCTGGTGCCTTTTTCTGTCCCCCCGATGGAGATTATCTGCATG
TCGTGGTCGTGTCCTGATCGAGTCGTCGTTGAGTCCCTATGTTTTTCTTCAAGAAATGTG
AGTCCTATGTCAGTCTGGTTGCGTTTGTGAACATTTCTGCGGCTGAGCAGCAGTTTGGTT
GGAACTGTGCAATGAAATAAATTGAACCCTGGTTTCTGGTTAAAAAAAAAAA

SCLO 3_a

GCACACCGTGAAACCAACCACTTTATTTATTACTGATAATACTCGCAGAAATATATACATG
TACCGTGCAGGCCATAAGGCACGGGCAACGATCGACGACCTGCAACTCCGAATACCGTT
CGGAGTATTTATTCTTCACACATTTATTTATTATTAGGCACAGACGGAACAGAAGTTATAG
TTGGTCTGGCGTCCATCACGAACATGGTACCCGTGGAAAGTTTGCACGCATTAAGCTCC
CGTCGTAGTCCACACTTGTCGATCAAGCGTCGAAGTCCAAGCGCCACAGGTTAATCCCA
CAGTTGGTAGAAGCGCCCAGGGGGGTATAGAAGGCATGTATGTTGATCTTATACATCCT
CCCGCCATTGGTTTGGTGCACAACCTTATGGTTCACAACACCGCG

SCLO 4_a

CGGCGAGCTCTATCGCTACATAGCTGATACTCAtGGTGCTTTTGTACATWHWTTNTTTTA
TGAAGCATTTGGTCTCACTGTCGTGGAGGCCATGACCTGTGGACTTCCTACTTTTGCAAC
ACTCCATGGAGGACCAGCTGAGATTATAGAGCATGGCATCTCGGGCTTCCACATTGACC
CGTACCACCTCGAGCAGGCTACTAATTTGACGGCTGACTTCTTCGAGCGATGCAAGCAA
GACCCAAATAACTGGGTGAAAATATCTGAAGCAGGACTGCAGCGCATATACGAAAAGTA
CACATGGAAGATATACTCTGAGAGGtTGATGACATTGGCCGGGGTC

SCLO 4_b

AACCAACCCATCGATCCAATTGTTCACTCGCTCAGTCGCTCGATCTCCCTCCAACAAGGT
CATCTCTCTCTTCTCGTCTCCTCATTACCCATGGCGGAGGAGAAGCACCACCACCACCT
GTTCCACCACAAGAAGGACGAGGAGGAGCAGGTGGAGCAGCCCGCCGGCGGCGGCGG
GTACGGCGAGGCCGCCGAGTACACGGAGACCACGGTGACGGAGGTGGTGTCCACGGG
CGAGGACGAGTACGACAAGTACAAGAAGGAGGAGAAGGAGCACAAGCACAAGCAGCAC
CTCGGCGAGGCCGGCGCCATCGCCGCCGGCGCCTTCGCACTCTACGAGAAGCACGAG
GCGAAGAAGGACCCGGAGCACGCGCACCGCCACAAGATCGAGGAGGAGGTCGCGGCG
G

SCLO 4_c

GGACAACTTACGTTGGCACGTGGTGTTGTGAACCACAAGGTTGTCCGCCAAACTAGTGG
cTGGAGGATGTATAAGATCAACATACAAGCCTTCTATACCCCCCTGGGCGCTTCTACCAA
CTGTGGGATTAACCTGTGGCGCTTGGACTTCGACGCTTGATCGACAAGTGTGGACTACG
ACGGGAGCTTAATGCATGCAAACTTTCCACGGGAACCATGTTCGTGATGGACGCCAGAC
CAACCATAACTTCTGTTCCGTCTGTACCTAATAATAAATAAATGTGTAAAGAATAAATACT
CCGAACGGTATTCGGAGTTGCAGGTCGTCGATCGTTGCCCGTGCCTTATGGCCTGCACG
GTACATGTACATATTTCTGCGAGTATTATCAGTA

SCLO 5_C

CGCGCACCTCTCCGCCGAGGACGAGGCGACGCTGGAGGCCGAGGTGCGGGGGTACTA
CGACGAGGCGGCGCCCAAGCGCCACACCAAGCCCTCCCGCAGCGAGCACTCCGCCGT
GTACGCCGACGCGCTCGTCCCGGACGCCGGCGGCAACACCCATCCGGAGCTCGACAA
GTTCCAAGAGCTCGAAGCACACACCGAGAGGTTGGCGTACGAGAGCGGCAATGTCGGA
GAGGAGTCCGTGGAGACGGAGTACTACAAGGATCTCGGCGGCGTTGGCAAGCAGCACC
ACACGACCGGAACGGGCTTCATCAAGATTGACAAAGCTAAGGGCGCCTCGTTCTAATTG
TCTGAAGACCCCAGCGCAGAGGAGCGCCATGCTTCTTGCAAGGGAAACCCTGCTACCAA
TGAGTGGATCCS

SCLO 6_a

GTGAACCACAAGGTTGTGCGCCAAACCAATGGCGGGAGGATGTATAAGATCAACATACA
TGCCTTCTATACTCCCCTGGGCGCTTCTAGCAACTGTGGGATTAACCTCTGGCGCTTGG
ACTTCGACGCTTGATCGACTAGCGCGGACTACAACATGAGGACCATGTTCTTCGTCGAC
GCTTAATGCATGGAAACTTCCCACGGGGGACCGTGTTCTTCATGGACGCCAGACCAACC
ATAATTTCTTTTCCGTTTGTACTGTCAACAAATATAAATATGTAAAGCATAAATCCGAACTG
TATTCGGAGTTGCAGGTCGTCGTTGCCCGTGCCTTATGGCCTGCACGGTACATGTACAT

ATTTCTGTCAAGTTCTGCGAGTATTTTAAGTAATAAATAAAGTGGTTGGTTTCAAAAAAAA
AAAAAAAAAAAAAAAAAAAAAAAA

SCLO 7_b

CAAGAACATTAGCTTCACCGTCTGGGATGTCGGGGGTCAGGACAAGATCAGACCTCTCT
GGAGGCATTACTTCCAGGACACCCAGGGTCTTATCTTTGTTGTGGACAGCAATGACCGT
GATCGTGTTGTTGAAGCAAGAGATGAGCTCCACAGGATGCTGAACGAGGATGAGCTACG
TGATGCTGTGCTGCTTGTTTTTGCCAACAAGCAAGATCCTCCCAATGCCATGAATGCTGC
TGAGATTACTGACAAGCTTGGATTACACTCCCTGCGCCAGCAACACTGGTACATCCAGA
GCACCTGTGCTACAACTGGTGAGGGTCTCTATGAGGGCCTGGACTGGCTGTCCAGCAA
CATTGCGAGCAAGGCTTGAGTCCTGCCTTGGATA

SCLO 8_a

CGGCTTGAAATTGGCAACGAAAAAAAACCTATATGCGGGCCTTGTTTAGATTGAAAGTTTT
TTCAACCCGATGAATAGTACCACTTACGCCTTATTTGACAAATATTGTCAAATCGAGGACC
AACTAGGCTCAAAAGGTTCATCTCGTGATTTCC

SCLO 8_b

GAGATGGCAAATGCAAGATCGGGTGTCGCCGTGAATGACGAGTGCATGCTGAAGTTCG
GCGAGCTTCAGTCGAAGAGGCTGCACCGCTTCATAACTTACAAGATGGACGACAAGTTC
AAGGAGATAGTTGTGGACCAGGTTGGGGATCGCGCCACCAGCTACGAGGACTTCACAA
ACAGCCTCCCTGAGAATGACTGCCGGTACGCAATCTATGATTTCGACTTTGTGACTGCAG
AGGATGTCCAGAAGAGCAGGATCTTCTATATCCTATGGTCCCCAGACTCCGCCAAGGTG
AAGAGCAAGATGCTTTACGCAAGCTCAAACCAGAAGTTCAAGAGTGGGCTCAATGGCGT
TCAGGT

SCLO 8_c

GCTTCTGGTGCCACTTCTTCCACTCGAAGTAGAGAGATTGAATGCCATAAGTGTCATGGC
TGTGGGCACATTGCTGACAATGTCCAAGTAGAAGGACCATGTTGTTGTATGAAAAGGGA
GAATGGGAATCTGAAAGTGATCCCGAGG

SCLO 9_a

CAAACATCACACATGCATTAGATTGAAACTACCAATCTAGTTAATCTCCAACATCACACAC
TACACTAACCAACTAATTTTACGTACGTCGGTACATGATGATACGCACTTACTATACATAC
ACTTCCCTCCTAGGAAAGGATTCTGATCCTGCTCGATCGCATTTCCTCTTGCCATGCTGC
TACCACGTCGCCCAGCTCCACCCGGAGCTCCTCCACCCGAACGCCTCGGCCTCGTCGC
CGTCGCCGCCCTCCAGCCTCAGCTCCACTAGCATCAGCTCGTGCGACACCTCCGACATC
TCCATCGACATGGCCACGAGACCCGACAGGCCGCCGAGCGGCAGGCGCTCGACTCCCT
TCTTCCCCACCATGAAGTAGCCGAGCTTGGCGCCAGC

SCLO 9_b

GTCTCTACTAACCTCAACAGTTCTGTAGCAAATCTGTACAAGCTTAGATAGATAGCTGATC
ACATTTTTCTTTCCCGTTTGATTTTTTTGAGGTTCGAATTTAGGAAATGGAGGCGGTTTGT
TCTGAACTAGTGATACGATTTGATGCTAGCGTTTTCTTGTGTGGTGGTCTGGTGGTTTGC
TTGTGAATTGAGAACCTTTTACAGTTTCAGATTATTAATTTTGATATTATCCACAAAAAAAA
AAAAAAAAAAAAAAA

SCLO 9_c

TTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTACGGCTATAACTAAATTCCTC
GATAAATATCAGTCACTATATCCAGGGTACAGATGGACATGAAACTAATAACAACATCAC
AATCAAAATTTTTAGACTCAAAAAAGTTTGTGGAGACTGGTGGTCCAAGTCACCGTGCTG

SCLO 10_a

GGGGTACCGTGCGCTTGTGGATTTATCTGTGTGCGTTAATTTATACCAACACTTGTGGAT
CCCCATCCTGTGTTGGTTGTGCGACACCTGATTGCGGGTATGGCATGTGATGCCAATTA
GCGTG

SCLO 11_a

ACCGAAAAAAGAGACCTTTTTCGCTCGGCGCCGCACTATGCTCCGCTTCAACAAGTGA
GAGTTTTCGGCGGAACCGGTGAACCACGCGAGCTGGTTAGATGCGTGGGGCAGAGCC
GATTTAGCCTCATAACCAGGCCTACCTGGGGAGTTCGGCAACATCGCCGTCCAGGACT
GGCCTGTGACCAAAGGGGATGGTAGCGACGCAACCGTCGTTGGACGTGCACAGGGCAT
CCAGTTCAAACCAAGCGAGAGGAACGACCAAGCCTGGTATACCACCTTGACCATAGTGT
TCGAGCGCACGAGCCTCAGGGGATCCACGCTTCAGATGATGGGTTACATCCCACAAGAT
GGTCAGTGGAGCATTTTTGGAGGAACTGGACAACTTACGATGGCACGCGGTGTTGTGAA
CCACAAGGTTGTGCGCCAAACCAATGGCGGGAGGATGTATAAGATCAACATACATGCCT
TCTATACTCCCCTGGGCGCTTCTAGCAACTGTGGGATTAACCTCTGGCGCTTGGACTTC
GACGCTTGATCGACTAGCGCGGACTACAACATGAGGACCGTGTTCTTCGTCGACGCTTA
ATGCATGGAAACTTCCCACGGGGGACCGTGTTCTTCACGGACGCCAGACCAACCATAAT
TTCTTTTCCGTTTGTACTGTCAACAAATATAAATATGTAAAGCATAAATCCGAACTGTATTC
GGAGTTGCAGGTCGTCGTTGCCCGTGCCTTATGGCCTGCACGGTACATGTTCATATTTC
TGTCGAGTTCTGCGAGTATTTTAAGTAATAAATAAAGTGGTTGGTTTCACGGTTAAAAAAA
A

SCLO 22_a

CGTGCATGCGTATGTCGGTACGTGCTTGTTTATGCAGTGTGATGGTATATGATATAGCTG
GTGTGCCAGATTCCCCTGACGAAGGCTGGCTGACCTGTATGGTCTGAGAAAATTGTTTA
GGTATCTAGAGATGTAAGAAAACATTCTTCAGTTCAGTTTCCCCGTGTGGAGT

SCLO 13_a

GAAAAACCGAACTGAAACTAATATAGAAACCCAACATAAGGGCAAGGTTGTGGCATGTCC
ACAAATACACCAAATAGAGAGGAAGAAAAACTACAACGCCAAGGCATATTAACTCTTCGA
TGATGCCTCCAAGGAGGTAAATGACGCCACTACGTAATCATCGTCGGTCAAGCAACAAG
TGGTGACGCTGAGCGCTTGGCAAGGCTTTCACCCGAGCCTTGT

SCLO 14_a

CTTCCTATATCCTACGATGGGTGATGACTAGGACCAAGGGCATGGTCACCTCATATGTCT
TACCTGATGCTTTTGTTGGACATGACTACTATTTGGCACTGTAATCCAAACTAAGTGTGTG
GCCTCGAAACTATTTGCCTATGAGGCTAA

SCLO 15_a

GAATTCCGATTTACGCCTCATAGGATCACAGAGTCAGGAGGTCGAGACCAGCCTGACCA
ACATGATGAAACCCTGTCTTTACTGAAAATACAAAAATTAGCCGAGCACAGTGGCTGCTG
TGCCCGTAATCCCCGCTACTCAAGAGGCTGAGGTAGAAGAATTACCTGAACTTGGGAGG
TGGGGGTTGCATATGAGGCTAAATCGGAATTCCGATTTAGGCTCATATTTTAAGCCAGAC
CACTTGAGTTGGGTGCCTGATAATACT

SCLO 16_a

CTTGTCATACCCACCATAGTTCTAAACCTTCCGCTATAACTGCCGCTATAGCCCGCTGTA
GCCTTTTGAGGCAGGGCGCCGCTATTTGAGTTCATGTATAATTTAGCCGTTATATTCCGC
GCTATTAGCTATTTTAGATATATACTGCTAAACGTCTTAGCCCGC